UNIVERSIDADE DE SÃO PAULO

FACULDADE DE CIÊNCIAS FARMACÊUTICAS DE RIBEIRÃO PRETO

Quimiometria aplicada à metabolômica de *Aldama* La Llave: Contribuições quimiotaxonômicas e fitoquímica direcionada baseada em inibição de cicloxigease-1 e 5-lipoxigease

Felipe Antunes dos Santos

Ribeirão Preto

2014

UNIVERSIDADE DE SÃO PAULO

FACULDADE DE CIÊNCIAS FARMACÊUTICAS DE RIBEIRÃO PRETO

Quimiometria aplicada à metabolômica de *Aldama* La Llave: Contribuições quimiotaxonômicas e fitoquímica direcionada baseada em inibição de cicloxigease-1 e 5-lipoxigease

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciências Farmacêuticas para obtenção do Título de Mestre em Ciências

Área de Concentração: Produtos Naturais e Sintéticos

Orientado(a): Felipe Antunes dos Santos

Orientador(a): Prof. Dr. Fernando Batista da Costa

Ribeirão Preto

2014

SANTOS, F.A.	**Quimiometria aplicada à metabolômica de *Aldama* La Llave: Contribuições quimiotaxonômicas e fitoquímica direcionada baseada em inibição de cicloxigease-1 e 5-lipoxigease**		MESTRADO FCFRPUSP 2014

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Santos, Felipe Antunes

Quimiometria aplicada à metabolômica de *Aldama* La Llave: Contribuições quimiotaxonômicas e fitoquímica direcionada baseada em inibição de cicloxigease-1 e 5-lipoxigease. Ribeirão Preto, 2014.

123 p.; 30cm.

Dissertação de Mestrado, apresentada à Faculdade de Ciências Farmacêuticas de Ribeirão Preto/USP – Área de concentração: Produtos Naturais e Sintéticos.

Orientador: Da Costa, Fernando Batista.

1. *Aldama* La Llave. 2. Quimiotaxonomia. 3. Fitoquímica Direcionada.

FOLHA DE APROVAÇÃO

Felipe Antunes dos Santos

Quimiometria aplicada à metabolômica de *Aldama* La Llave: Contribuições quimiotaxonômicas e fitoquímica direcionada baseada em inibição de cicloxigease-1 e 5-lipoxigease

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciências Farmacêuticas para obtenção do Título de Mestre em Ciências

Área de Concentração: Produtos Naturais e Sintéticos.

Orientador(a): Prof. Dr. Fernando Batista da Costa

Aprovado em:

Banca Examinadora

Prof. Dr. _____

Instituição: _____ Assinatura:_____

Prof. Dr. _____

Instituição: _____ Assinatura:_____

Prof. Dr. _____

Instituição: _____ Assinatura:_____

Aos amores da minha vida:

meus pais, José e Sebastiana;

meu irmão Fabrício;

e à Rose...

AGRADECIMENTOS

- Primeiramente agradeço à toda minha família pelo apoio incondicional, por todo o carinho, por não medirem esforços ao me ajudar, por me enxergarem melhor do que sou, por me incentivarem até o fim, por estarem sempre ao meu lado... Por tudo isso e tanto mais que não caberia aqui em palavras, eu me senti mais forte nos momentos difíceis e consegui concluir esta Dissertação;

- Segundo, agradeço infinitamente ao prof. Dr. Fernando Batista da Costa por ter acreditado e confiado no meu trabalho... Por ter me dado oportunidade. Grande orientador, conselheiro, mentor, e, acima de tudo, grande amigo... Tenho-o por admiração não apenas por ser meu orientador, mas por ser a pessoa boa que é...

- Agradeço a todos os amigos do laboratório de Farmacognosia: à Ingrid, grande colaboradora e amiga desde o primeiro ano de faculdade; ao Bruno, Tiago Branquinho, Ricardo, Rosana e Daniela Chagas, pelos ensinamentos e ajuda; à Amanda, Paula, Danni e Rejane pelas colaborações científicas; ao Lucas e à Anny, pelo companheirismo e apoio; ao Federico, pelas ricas conversas teóricas e por compartilhar conhecimento; ao Gari, pelo apoio decisivo ao final deste trabalho. Mais uma vez, agradeço a todos não somente pelo apoio científico, bem como pela amizade;

- A todos do laboratório de Farmacognosia, pelo apoio e contribuições científicas: prof. Jairo; profa Niege; Éric, Tati, Cris, William, Ju, Mauro, Marcela, João Luiz, Mariza, Eli, Jonas, Marivane, aos técnicos Mário e Angélica;
- Àqueles dos demais laboratórios: prof. Gobbo; prof. Norberto; prof. Vessecchi; prof. Renato Lajarim; prof. Marcus Scotti; profa. Mônica; prof. Sérgio Ambrósio; ao técnico Tomás e ao Lucas, aluno de doutorado do prof. Norberto;
- À Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo;
- À CAPES, CNPq e, principalmente, à FAPESP pelo apoio financeiro;

...Enquanto eu escrevia estes agradecimentos, fiz uma pausa e conversei com meu cachorro (que às vezes trato como filho), mesmo com ele não entendendo qualquer coisa do que digo... Pensei que, se cães soubessem ler, não somente os agradecimentos, mas também a dedicatória iria a eles... Um cientista, em virtude de ceticismos e curiosidades à cerca do funcionamento da natureza, inerentes e necessárias à sua profissão, às vezes é tomado a se perguntar onde estaria Deus em sua obra. Onde ele está e o que é ele? Teria ele criado as leis da física, a considerado tão perfeita e harmoniosa que não queira interferir nelas? Seria ele a própria física (e metafísica) que rege este Universo? Poderia ele ler os meus agradecimentos ou pensamentos? E se pudesse, iria se dispor a isto? É certo que, por hora, não encontrarei estas respostas... e nem sei se sou digno delas... De modo que, pondo ao lado a paixão humana em querer conhecer e aprender, tento-me colocar em posição mais humilde, supondo que, de algum modo, Ele tenha feito parte desta minha vitória: Agradeço a Deus por eu ter concluído esta Dissertação.

"*Não importa quanto a vida possa ser ruim, sempre existe algo que você pode fazer, e triunfar. Enquanto há vida, há esperança*"

Stephen Hawking

RESUMO

SANTOS, F. A. **Quimiometria aplicada à metabolômica de *Aldama* La Llave: Contribuições quimiotaxonômicas e fitoquímica direcionada baseada em inibição de cicloxigease-1 e 5-lipoxigease.** 2014. 143f. Dissertação (Mestrado). Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2014.

As espécies do gênero *Viguiera* Kunth, recentemente transferidas para *Aldama* La Llave (Asteraceae, Heliantheae), ainda possuem problemas de delimitação taxonômica. Tal gênero possui 35 espécies distribuídas por todo o território brasileiro, principalmente no cerrado. Análises químicas têm demonstrado seu potencial em auxiliar na resolução de problemas em vários níveis taxonômicos com base em grupos químicos de metabólitos secundários. Além disso, análises fitoquímicas tem revelado que determinadas substâncias de *Aldama* possuem potencial para serem estudadas tendo vista a investigação de mecanismos moleculares anti-inflamatórios. Desta forma, foi proposto neste trabalho realizar a abordagem da metabolômica não-direcionada auxiliada por quimiometria, visando-se fornecer dados químicos tanto para contribuições quimiotaxonômicas quanto para encontrar substâncias bioativas. Por meio de tal abordagem, análises multivariadas não supervisionadas (PCA e HCA), bem como supervisionadas (OPLS-DA), com dados provenientes de UHPLC-DAD-Orbitrap, demonstraram que o gênero *Aldama* é quimioconsistente e, deste modo, determinadas substâncias químicas discriminantes foram sugeridas. Além disso, espécies que apresentam problemas taxonômicos tiveram os seus posicionamentos infragenéricos explicado do ponto de vista quimiotaxonômico. Quanto às análises para encontrar substâncias bioativas (fitoquímica direcionada), os alvos inflamatórios pesquisados foram as enzimas pró-inflamatórias COX-1 e 5-LOX, com as quais se realizou ensaios *in vitro* verificando-se a atividade positiva de extratos de espécies de *Aldama*. *A. trichophylla* foi a espécie selecionada para realizar a fitoquímica direcionada, uma vez que apresentou inibição dupla contra as enzimas e por se tratar de uma espécie muito pouco estudada do ponto de vista fitoquímico. Os melhores modelos de inibição de COX-1 e 5-LOX criados com dados químicos de *Aldama* foram selecionados a partir de diversas combinações de algoritmos, *softwares* de processamento (*MZmine*, *MetAlign* e *MSClust*) e técnicas-hifenadas, tais como UHPLC-Orbitrap e HPLC-TOF. Deste modo, determinados picos de íons foram apontados pela quimiometria como sendo discriminantes para a atividade de inibição dupla. Tais picos foram desreplicados com base em seus perfis de UV e padrões de fragmentação via HCD-Orbitrap e Ion-Trap. Deste modo, foi possível desreplicar três prováveis substâncias químicas inibidoras de COX-1 e 5-LOX: kaempferol-3-*O*-glucuronideo, quercetina-3-*O*-metil-7-glucuronideo e kaempferol-3-*O*-(6"-malonil-glucosideo). Por fim, foi realizada a fitoquímica direcionada. O fracionamento cromatográfico permitiu isolar tais substâncias e uma análise preliminar de RMN ^{1}H foi realizada com o intuito de realizar a identificação estrutural. Tais substâncias terão a suas atividades confirmadas na inibição das duas enzimas em um futuro próximo.

Palavras-chave: *Aldama* La Llave, Quimiotaxonomia, Fitoquímica direcionada, Quimiometria, Metabolômica não-direcionada

ABSTRACT

SANTOS, F. A. **Chemometrics applied to metabolomics of *Aldama* La Llave: chemotaxonomic contributions and targeted phytochemistry based on cyclooxygenase-1 and 5-lipoxygenase inhibition**. 2014. 143f. Dissertation (Master). Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2014.

The species of the genus *Viguiera* Kunth, recently transferred to the genus *Aldama* La Llave (Asteraceae, Heliantheae), still show problems of taxonomic boundaries. This genus has 35 species distributed throughout the Brazilian territory, mostly in the cerrado. Chemical analyses has demonstrated its potential to help in the solution of problems at various taxonomic levels based on chemical groups of secondary metabolites. Furthermore, phytochemical analyses have shown that certain compounds from *Aldama* have the potential to be studied aiming to investigate anti-inflammatory molecular mechanisms. Thus, in this work it was proposed to carry out the untargeted metabolomic approach aided by chemometrics, aiming to provide chemical data either for chemotaxonomic contributions and to find bioactive compounds. Through this approach, both unsupervised (PCA and HCA) and supervised multivariate analysis (OPLS-DA) combined with data from UHPLC-DAD-Orbitrap analyses showed that the genus *Aldama* is “chemoconsistent” and thus certain discriminant compounds were suggested. Additionally, some species with taxonomic problems had their infrageneric positioning explained from the chemotaxonomic point of view of. With regards to the analyses carried out to find bioactive compounds (targeted phytochemistry), the inflammatory targets investigated in this study were the COX-1 and LOX-5 pro-inflammatory enzymes with which *in vitro* assays were made and positive activity of extracts from species of *Aldama* was observed. *A. trichophylla* was the selected species for the targeted phytochemistry because it showed dual inhibition activity against both enzymes and it is still little investigated from the chemical point of view. The best models for inhibition of COX-1 and 5-LOX obtained with chemical data of *Aldama* were selected by means of various combinations of algorithms, processing software (MZmine, MetAlign and MSClust) and hyphenated techniques, such as HPLC-TOF and UHPLC-Orbitrap. Thus, certain peak ions were appointed by chemometrics as discriminant for the dual inhibition activity. These peaks were dereplicated based on their UV profiles and fragmentation patterns via HCD-Orbitrap and Ion-Trap. Thus, it was possible to dereplicate three probable chemical inhibitors of COX-1 and 5-LOX: kaempferol-3-*O*-glucuronide, quercetin-3-*O*-methyl-7-glucuronide and kaempferol-3-*O*-(6"-malonyl-glucoside). Finally, the targeted phytochemistry was carried out. The chromatographic fractionation allowed us to isolate these three compounds and a preliminary ^{1}H NMR analysis was performed in order to conclude their structural identification. In the near future, these substances will have their activity evaluated by the inhibition of the two enzymes.

Keywords: *Aldama* La Llave, Chemotaxonomy, Targeted Phytochemistry, Chemometrics, Untargeted Metabolomics

LISTA DE FIGURAS

Figura 1 - Diterpeno *ent*-caurano e LST furanoeliangolido orientado em β, comumente encontrados em Heliantheae.....3

Figura 2 - Níveis de especiação de esqueleto carbônico de LST.....4

Figura 3 - Tipos de análises metabolômicas (Retirado de Krastanov, 2010; pag 1538).....8

Figura 4 - Comparação de diferentes softwares de processameto de dados (Extraído de Castillo *et al*., 2011).....11

Figura 5 - População de *A. trichophylla*.....14

Figura 6.a - Tricoma Gladular de *A. robusta*.....16

Figura 6.b - Tricoma Glandular pinçada por agulha.....16

Figura 7.a - Antera de inflorescências de *A. arenaria*.....16

Figura 7.b - Tricomas gladulares dem anteras de *A. arenaria*.....16

Figura 8 - Exemplo do *Pareto scaling* no tratamento de dados químicos provenientes de EM.....23

Figura 9.a - Sephadex.....26

Figura 9.b - Coletor automático.....26

Figura 10 - HCA realizada com dados químicos provenientes dos modos de ionização positivo e negativo do ET de *Aldama* La Llave (em vermelho) e todos os outros gêneros de Asteraceae.....29

Figura 11 - HCA realizada com dados químicos provenientes do modo de ionização positivo do ET de *Aldama* La Llave (em azul) e todos os outros gêneros de Asteraceae.....30

Figura 12 - HCA realizada com dados químicos provenientes do modo de ionização negativo do ET de *Aldama* La Llave e todos os outros gêneros de Asteraceae.....31

Figura 13 - PCA com dados provenientes de ambos os modos de ionização de ET de *Aldama* La Llave e outros gêneros.....33

Figura 14 - Classificação de Blake (1918) (Extraído de Da Costa, 2008 – Texto Sistematizado de Livre Docência).....34

Figura 15 - HCA com base nos dois modos de ionização, ampliada, para o gênero *Aldama* La Llave (em vermelho).....35

Figura 16 - PCA com base em ambos modos de ionização para o gênero *Aldama* La Llave, de acordo com a classificação infragenérica de Blake (1918) e o *outlier A. aspillioides*.....38

Figura 17 - PCA com base em ambos modos de ionização para o gênero *Aldama* La Llave, de acordo com a classificação infragenérica de Blake (1918).....38

Figura 18 - *Overfitting* na PCA, mostrando o caso em que o ajuste melhora, mas a predição piora.....39

Figura 19 – OPLS-DA com base na classificação de Blake (1918) realizado com ambos os modos de ionização para *Aldama* La Llave.....40

Figura 20 - Ilustração da remoção de variáveis não importantes abaixo de 1,0.....41

Figura 21 - OPLS-DA otimizada com VIP > 1,0 realizado com base na Classificação de Blake e dados adquiridos por ambos modos de ionização para *Aldama* La Llave.....42

Figura 22 - Valores de ajuste e predição alcançados pelo modelo OPLS-DA otimizado (VIP > 1).....43

Figura 23 - Valores de predição das três espécies não classificadas por Blake.....43

Figura 24 - Classificação climática no Brasil (Extraído de Alvares, 2014).....45

Figura 25 - OPLS-DA com base na classificação climática de Alvares (2014) realizada com dados químicos de *Aldama* La Llave.....46

Figura 26 - PCA realizada com picos de alta intensidade detectados por ambos modos de ionização..48

Figura 27 - HCA realizada com picos de alta intensidade detectados por ambos modos de ionização.....49

Figura 28 - OPLS-DA realizada com dados provenientes do modo positivo e negativo para o gênero *Aldama* La Llave e outros gêneros.....49

Figura 29 - Estruturas químicas de algumas das prováveis substâncias químicas do gênero *Aldama* La Llave desreplicadas.....51

Figura 30 - OPLS-DA para modelo de predição de tipos de extração.....55

Figura 31 - Valores de ajuste alcançados para dados provenientes do TOF e tratados por diferentes algoritmos do *MZmine*.....57

Figura 32 - Valores de ajuste alcançados para dados provenientes do OBT e tratados por diferentes algoritmos do *MZmine*.....57

Figura 33 - Indicação de *overfitting* (Extraída do Manual do SIMCA 13.0.3).....61

Figura 34 - Modelo de inibição enzimática por OPLS-DA criado com dados provenientes do modo negativo do Orbitrap e processados pelo MZmine (3 neg OBT).....63

Figura 35 - RMSEcv em decaimento com o aumento do número de componentes do modelo de predição.....64

Figura 36 - Destaque para o processo que gera o *Contribution Plot* para *A. trichophylla*.....65

Figura 37 - Variáveis geradas pelo *Contribution Plot* para *A. trichophylla*.....65

Figura 38 - Variáveis geradas pelo *Coefficient Plot* para inibição de COX-1 e 5-LOX.....66

Figura 39.a - Substância discriminante para o modelo "MZmine 4 TOF" apontada pela Decision Tree.....68

Figura 39.b - Substância discriminante para o modelo "MZmine + MSClust TOF 4 neg.....68

Figura 40 - Gráfico de *loadings* evidenciando a substâncias de código 545 (*m*/*z* 523) para o modelo "MZmine TOF 4 neg".....68

Figura 41 - Gráfico de *loadings* evidenciando a substâncias de *m*/*z* 523 (código 545) para o modelo "MZmine TOF 4 neg".....69

Figura 42 - Esquema de compartimentos do Orbitrap Exactive Plus (Extraído de planetorbitrap.com - Thermo Fisher).....70

Figura 43 - Perfil químico do espectro de ultravioleta registrado em 254 nm para ET de *A. trichophylla*.....71

Figura 44 - Perfil químico do espectro de ultravioleta registrado em 254 nm para cafeína e flavona, respectivamente.....72

Figura 45 - Perfil químico registrado em 254 nm para *A. trichophylla*.....72

Figura 46 - Perfil químico registrado em 325 nm para *A. trichophylla*.....73

Figura 47 - Perfil de UV típico de ácido cafeoilquínico.....73

Figura 48 - Espectro de massas do ácido 1,4-di-*O*-cafeoilquínico gerado no modo de ionização positivo.....74

Figura 49 - Espectro de fragmentação do ácido 1,4-di-*O*-cafeoilquínico gerado pelo HCD no modo positivo.....74

Figura 50 - Espectro de fragmentação do ácido 1,4-di-O-cafeoilquínico gerado pelo HCD no modo negativo.....74

Figura 51 - Esquema de movimento dos íons no Orbitrap (Extraída de Hu *et al*., 2005).....76

Figura 52 - Perfil químico da fração AcOEt registrado em 254 nm para *A. trichophylla no* Ion-Trap.....77

Figura 53 - Perfil de UV para kaempferol-3-*O*-glucuronideo.....78

Figura 54 - Espectro MS^2 do kaempferol-3-*O*-glucuronideo, evidenciando o fragmento 287 gerado a partir da molécula protonada de *m*/*z* 463.....78

Figura 55 - Espectro MS^3 do kaempferol-3-O-glucuronideo, evidenciando fragmentos que se formam a partir do íon de *m*/*z* 287, o qual foi gerado a partir da molécula protonada de *m*/*z* 463.....78

Figura 56 - Modelo “MZmine Orbitrap 3 neg” evidenciando as substâncias quercetina-3-*O*-metil-7-glucuronideo (*m*/*z* 491) e kaempferol-3-*O*-glucuronideo (*m*/*z* 461).....91

Figura 57 - Gráfico de *loadings* do modelo "MZmine TOF 1 pos" evidenciando a *m*/*z* substâncias de *A. trichophylla* discriminantes para a inibição de COX-1 e 5-LOX.....92

Figura 58 - Gráfico de *loadings* do modelo "MZmine TOF 1 pos" evidenciando o nome (em inglês) das substâncias de *A. trichophylla* discriminantes para a inibição de COX-1 e 5-LOX e tamarixetin-3-O-β-(6”-malonil-glucopiranosideo).....92

Figura 59 - Comparação estrutural de inibidores *dual* (presentes em Charlier e Michaux, 2003) e do flavonóide glicosilado.....93

Figura 60 - CCD analítica das frações de *sephadex*.....94

Figura 61 - Registro de algumas frações em 254 e 325 nm.....95

Figura 62 - EIC *m*/*z* 463 fora encontrada em 22,5 min no cromatograma do íon extraído para a fração S-16.....95

Figura 63 - Perfil de UV evidenciando se tratar da substância kaempferol-3-*O*-glucuronideo.....96

Figura 64 - Espectro de massas para a substância kaempferol-3-*O*-glucuronideo.....96

Figura 65 - Espectro de RMN ^{1}H apresentado para o pico 355.....98

Figura 66 - Valores de deslocamento preditos para os hidrogênios do ácido 1-O-*E*-cafeoilquínico.....99

Figura 67 - Numeração dos átomos de hidrogênio e seus respectivos valores de deslocamento químico dos ácidos 1-O-*E*-cafeoilquínicos.....99

Figura 68 - Espectro de RMN ^{1}H apresentado para o pico 535.....100

Figura 69 - Numeração dos átomos de hidrogênio e os respectivos valores de deslocamento químico do kaempferol-3-*O*-(6"-malonil-glucosideo).....101

Figura 70 - Numeração dos átomos de hidrogênio e seus respectivos valores de deslocamento químico para do kaempferol-3-O-(6"-malonil-glucosideo).....101

LISTA DE TABELAS, QUADROS, ESQUEMAS, GRÁFICOS E MECANISMOS

Tabela 1 - Lista de prováveis substâncias químicas para *Aldama* La Llave.....52

Tabela 2 - Resultado de inibição das enzimas COX-1 e 5-LOX para *Aldama* La Llave.....56

Tabela 3 - Prováveis substâncias (nomes em inglês) inibidoras de COX-1 e 5-LOX desreplicadas em alguns dos melhores modelos de predição criados via OPLS-DA.....81

Tabela 4 - Todas as substâncias desreplicadas (nomes em inglês) e seus respectivos padrões de fragmentação encontrado nos modos de ionização positivo e negativo via Ion-Trap e Orbitrap (em destaque as prováveis substâncias ativas contra COX-1 e 5-LOX).....82

Quadro 1 - Classificação infragenérica de Blake (1918).....36

Quadro 2 - Modelos otimizados com o máximo número de componentes ortogonais permitido.....62

Quadro 3 - Comparação dos valores de RMSEcv gerados pela *Decision Tree* (Weka) e OPLS-DA para os melhores modelos escolhidos.....67

Quadro 4 - Variáveis apontadas pela *Decision Tree*.....69

Quadro 5 - Informação sobre os valores de acurácia alcançados pelos analisadores TOF, Ion-Trap e Orbitrap.....90

Esquema 1 - Esquema geral da obtenção e processamento dos dados químicos dos extratos vegetais de *Aldama* em software de processamento de dados de LC-MS oriundos do TOF e do Orbitrap.....19

Esquema 2 - Diferentes algoritmos utilizados no software MZmine 2.9. pos: ionização no modo positivo; neg*: indica que o mesmo esquema foi utilizado para os dados provenientes para do modo negativo.....20

Esquema 3 - Partições líquido-líquido do ET.....25

Gráfico 1 - Capacidade preditiva de cada modelo criado com dados tratados via *MZmine*....58

Gráfico 2 - Comparação dos diferentes softwares de tratamento utilizados para dados provenientes do modo positivo e negativo do TOF e OBT.....59

Gráfico 3 - Melhores combinações para todos os modelos criados com base em dados químicos provenientes dos modos positivo e negativo do TOF e OBT.....60

Gráfico 4 - As melhores combinações otimizadas por meio da deleção de variáveis não importantes para a projeção (VIP < 1,0).....61

Mecanismo 1 - Fragmentação de flanonol do tipo quercetina. Fragmentos observados para isoquercetina.....86

Mecanismo 2 - Fragmentação para flavonol do tipo kaempferol.....87

Mecanismo 3 - Fragmentação da isorhammetima-3-*O*-glicosideo.....87

Mecanismo 4 - kaempferol-3-*O*-(6''-malonil-glucosideo).....88

LISTA DE ABREVIATURAS E SIGLAS

2D e 3D	bidimensional e tridimensional
COX	cicloxigenase
d	dubleto
DAD	*Diode Array Detector*
dd	duplo dubleto
DO	densidade ótica
EIC	*Extraction Ion Chromatogram*
ELF	extrato de lavagem foliar
ELI	extrato de lavagem de inflorescências
EM	Espectrometria de Massas
ESI	*electrospray ionization*
ET	extrato total
EtOH	etanol
GT	*Glandular Trichomes*
HAc	ácido acético
HCA	*Hierarchical Clustering Analysis*
HCD	*Higher-energy Collisional Dissociation*
HMBC	*Heteronuclear Multiple Bond Coherence*
HMQC	*Heteronuclear Multiple Quantum Coherence*
Hz	Hertz
IC	*Inhibitory Concentration*
J	constante de acoplamento
LC-MS	*Liquid Chromatography coupled to Mass Spectrometry* (designação geral)
LOX	lipoxigenase
LST	lactonas sesquiterpênicas
m/*z*	relação massa carga
MeCN	acetonitrila
MeOD	metanol deuterado
MeOH	metanol
MS	*Mass Spectrometry*
MS-MS	*Mass Spectrometry* - *Mass Spectrometry* (espectro de fragmentação)
NCE	*Normalized Collision Energy*
neg	modo de ionização negativo
OBT	Orbitrap
OPLS-DA	*Orthogonal Partial Least Square - Discriminant Analysis*
PCA	*Principal Component Analysis*
PG	prostaglandinas
pos	modo de ionização positivo
RMN	Ressonância Magnética Nuclear
RMSEcv	*Root Mean Squared Error - Cross Validation*

rpm	rotações por minuto
S	fração de *sephadex*
s	singleto
STL	*sesquiterpene lactones*
TG	tricomas glandulares
TIC	*Total Ion Chromatogram*
TOF	*Time-of-Flight*
VIP	*Variable Importance in Projection*

SUMÁRIO

RESUMO ____ **ix**

ABSTRACT ____ **x**

LISTA DE FIGURAS ____ **xi**

LISTA DE TABELAS e QUADROS ____ **xv**

LISTA DE ABREVIATURAS E SIGLAS ____ **xvii**

SUMÁRIO ____ **i**

Introdução ____ **1**

Objetivos ____ **13**

Material e Métodos ____ **14**

1 – Coleta e Secagem ____ *14*

2 – Extrato Total (ET) ____ *15*

3 – Coleta de Tricomas Glandulares (TG) ____ *15*

4 – Extrato de Lavagem Foliar (ELF) ____ *15*

4 – Extrato de Lavagem Foliar (ELF) ____ *16*

5 – Obtenção dos dados químicos via HPLC-PDA-TOF e UHPLC-DAD-Orbitrap™ e processamento dos dados via MZmine, MetAlign e MSClust ____ *17*

6 – Análise in vitro *do potencial anti-inflamatório dos extratos totais (ET)* ____ *20*

7 – Quimiometria dos dados químicos provenientes de LC-DAD-MS ____ *22*

8 – Desreplicação direcionada (Targeted dereplication) ____ *23*

9 – Isolamento dos picos alvos (Targeted phytochemistry) ____ *24*

10 – Elucidação estrutural via RMN ____ *26*

Resultados e Discussão dos estudos quimiotaxonômicos ____ **27**

1 – Hierarchical Clustering Analysis (HCA) e Principal Component Analysis (PCA) ____ *27*

2 – Orthogonal Partial Least Squares – Discriminant Analysis (OPLS-DA) e desreplicação dos metabólitos discriminantes ____ *37*

Resultados e Discussão dos estudos de inibição dupla das enzimas COX-1 e 5-LOX **55**

3 – Resultados dos ensaios de inibição enzimática COX-1 e 5-LOX & Comparação dos diferentes modelos criados ____ *55*

*4 – Otimização dos melhores modelos de inibição enzimática & Desreplicação direcionada*_____*60*
*5 – Fitoquímica direcionada das prováveis substâncias capazes de inibir COX-1 e 5-LOX*_____*93*
6 – Ressonância Magnética Nuclear de Hidrogênio (RMN ^{1}H)_____97
Conclusão_____ **103**
Referências _____ **105**

Introdução

A família Asteraceae (Compositae) é uma das mais numerosas dentre as Angiospermas, tendo cerca de 1.100 gêneros e 25.000 espécies (Funk *et al*., 2009). De acordo com Verdi e seus colaboradores (2005), por volta de "98% dos gêneros são constituídos de plantas de pequeno porte, sendo encontradas em todos os habitats, principalmente nas regiões tropicais da América do Sul". O gênero *Viguiera* Kunth (tribo Heliantheae, subtribo Helianthinae) possui 35 espécies brasileiras distribuídas em diversas regiões do território nacional, principalmente no cerrado (Magenta e Pirani, 2014; Blake, 1918). Tal gênero possui 180 espécies e apresenta problemas de delimitação taxonômica e têm sido amplamente estudado por botânicos, taxonomistas e fitoquímicos. Pesquisas recentes sugeriram uma posição derivada do gênero dentro da subtribo Helianthinae e uma origem recente dentro do bioma cerrado, bem como a necessidade de se realizar uma reavaliação taxonômica das espécies brasileiras (Schilling *et al*., 2000 e 2012).

Assim sendo, em 2006 foi realizada uma revisão taxonômica das espécies de *Viguiera* que ocorrem no Brasil (Magenta, 2006), verificando-se extensivamente a morfoanatomia do grupo, concluindo-se que havia a necessidade de se realizar estudos moleculares (com material genético) e fitoquímicos integrados às características morfoanatômicas do gênero. Por exemplo, as espécies *V. anchusifolia* e *V. megapotamica* apresentam morfologia externa muitas vezes semelhante, o que dificulta a sua delimitação. Além disso, foi observado que além das espécies *V. robusta* e *V. arenaria* serem facilmente confundidas em herbários e, muitas vezes, só distinguidas pelas características do grão de pólen, as espécies *V. discolor*, *V. squalida* e *V. bakerina* podem ser a mesma. Há outros problemas de delimitação que foram destacados por outros autores, tais como a difícil delimitação entre *V. linearifolia*, *V. filifolia* e *V. trichophylla*, por exemplo, sendo que as duas primeiras foram confundidas entre si por Blake (1884) e Blake (1918), ao passo que *V. trichophylla* foi considerada coespecífica de *V. linearifolia* por Robinson e Moore (2004). Não obstante todos estes problemas observados, recentemente os resultados obtidos das análises de ITS (região do espaçador ribossomal interno) com sítios de restrição de cpDNA e de combinações entre as análises de ITS e ETS (região do espaçador ribossomal externo), o grupo parafilético *Viguiera* foi reduzido a apenas uma espécie, *Viguiera dentata*, sendo, portanto, considerado agora um grupo monoespecífico, transferindo-se todas as outras espécies para o gênero *Aldama* (Schilling e Panero, 2002 e 2011). Entretanto, há de se levar em consideração que poucas espécies brasileiras foram

incluídas neste estudo. Ainda assim, Magenta considerou em seu último trabalho publicado o nome *Aldama* para as espécies brasileiras, destacando ainda novidades infraespecíficas e que havia um equívoco cometido por Blake ao diferenciar *V. bakeriana* de *V. subdentada*, sendo que ambas tratam-se da mesma espécie (Magenta, 2014). Portanto, entrando em perfeita consonância com os estudos morfoanatômicos e filogenéticos aqui apresentados, nesta dissertação o gênero *Viguiera* Kunth será de agora em diante designado como *Aldama* La Llave.

As últimas pesquisas têm demonstrado os avanços em se caracterizar e traçar uma delimitação taxonômica consistente para o gênero *Aldama*. Sendo assim, os caracteres químicos podem dar a sua contribuição às classificações taxonômicas naquilo que se denomina quimiotaxonomia. Em tempo, cabe aqui explicar *a priori* algumas definições conceituais antes de entrar nas contribuições químicas para a taxonomia de plantas. Primeiramente, é de suma importância mencionar que quimiotaxonomia não pode ser confundida com quimiossistemática. A primeira diz respeito à contribuição química para classificações taxonômicas, sendo que táxon significa uma unidade taxonômica, que por sua vez indica um determinado nível de sistema de classificação científica. Assim sendo, taxonomia é o estudo que descreve os taxa. Já quimiossistemática se trata da mesma utilização dos caracteres químicos para classificações taxonômicas, porém, integrados a outros estudos, como os filogenéticos, ambientais, sazonais, circadianos, ecoquímicos, etc. Fica claro que quimiotaxonomia é um estudo mais restrito, ao passo que quimiossistemática é um estudo mais amplo. Apenas para deixar mais claras as definições, pode-se mencionar aqui que filogenia é o estudo das relações evolutivas entre os seres vivos.

Deste modo, alguns químicos de produtos naturais têm utilizado a química em classificações taxonômicas, uma vez que têm se demonstrado que determinadas substâncias químicas do metabolismo secundário de plantas são específicas de determinadas famílias, tribos, gêneros, ou até mesmo de espécies. Por exemplo, para o nosso caso particular, pode-se citar os trabalhos de Seaman (1982) e de Zdero e Bohlman (1990) que evidenciaram que a família Asteraceae possui uma tendência em biossintetizar lactonas sesquiterpênicas (LST). Observaram também que a tribo Heliantheae tem uma tendência em biossintetizar diterpenos do tipo *ent*-caurano e LST do tipo furanoeliangolido com conFiguração β no carbono 8 (C-8), diferentemente das espécies de Vernonieae, cuja conFiguração no C-8 é sempre α. Pode-se visualizar estas estruturas químicas na Figura 1.

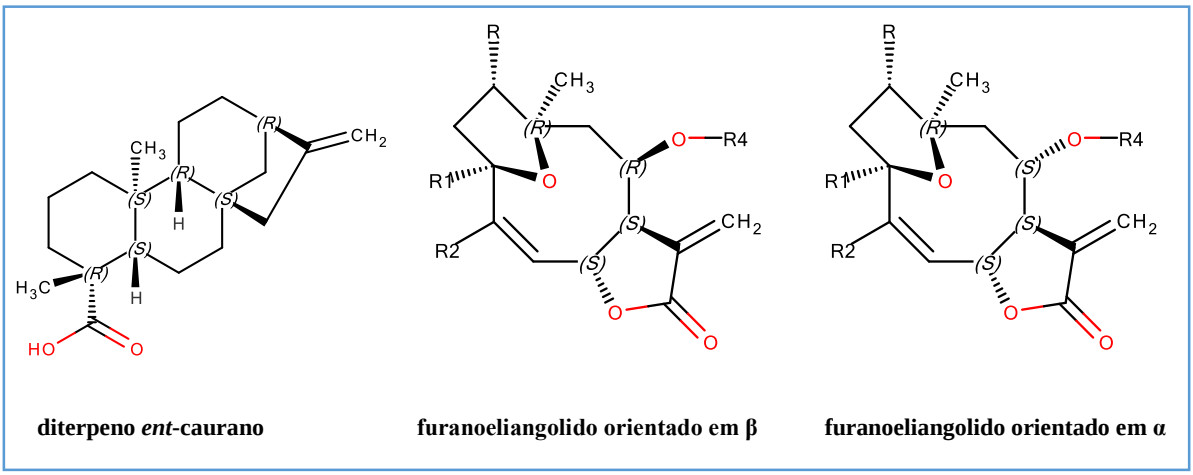


Figura 1 – Diterpeno do tipo *ent*-caurano e LST do tipo furanoeliangolido orientada em β no C-8, comumente encontradas em Heliantheae, e orientada em α no C-8, típicas de Vernonieae. Os grupos R, R1 e R2 podem ser átomos de hidrogênio ou hidroxila e R4 são grupos acila.

Também poderíamos notar que as LST do tipo furanoeliangolido 1-oxo-2-en-3,10-epoxieliangolido (budleína A – Figura não mostrada) não ocorrem nos gêneros *Tithonia* e *Helianthus* (Da Costa, 2008), os quais são filogeneticamente muito próximos a *Aldama*, sendo que esta é tida como um ancestral comum destes gêneros, segundo Blake (1918). Ainda assim, muitos outros metabólitos são compartilhados por estes gêneros e pode-se perceber que, especificamente em relação às LST, as classes químicas possuem esqueletos mais simples, tais como aqueles representados pelos níveis I e II da Figura 2 (Seaman, 1982). Com base nisso, pode-se também traçar vários níveis evolutivos para outras classes de metabólitos, tal como o exemplo das LST (Figura 2).

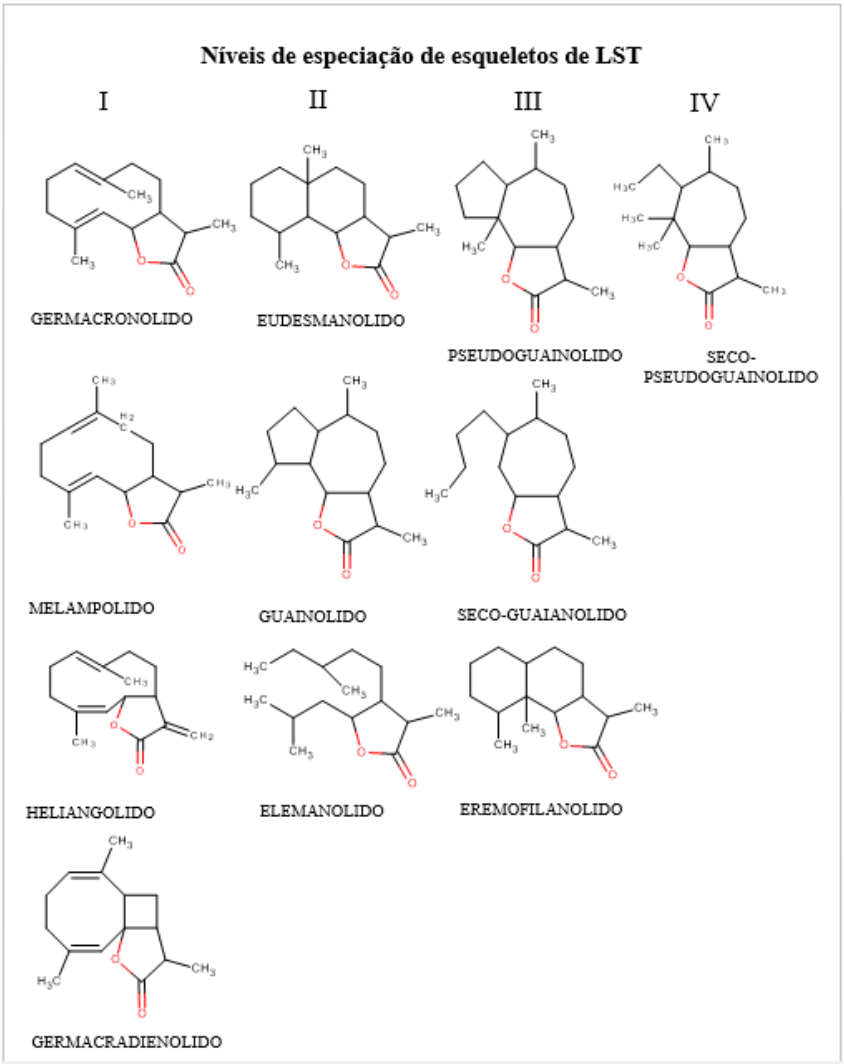


Figure 2 - Níveis de especiação de esqueleto carbônico de LST.

De fato, de maneira geral pode-se observar que níveis taxonômicos elevados de Asteraceae, tais como subfamílias e tribos, podem ser classificados de acordo com o tipo de esqueleto carbônico das estruturas químicas, ao passo que níveis mais baixos, como gêneros, espécies e variedades, podem ser classificados quimiotaxonomicamente de acordo com padrões de substituições, tais como grupos funcionais, estereoquímica e isomeria de posição, e níveis de oxidação (Da Costa, 2008).

Por falar em níveis de oxidação, cabe aqui mencionar a “teoria redox” da evolução química proposta por Gottlieb (1992) que descreve a relação diretamente proporcional entre os níveis de oxidação dos metabólitos secundários e o grau evolutivo das plantas a que eles pertencem. Esta teoria acredita ser o oxigênio um elemento que dá grau de especiação às estruturas químicas, possibilitando que elas participem de inúmeras reações enzimáticas, por

exemplo. Por fim, esta teoria postula que o elevado grau de especiação de metabólitos secundários confere não somente às plantas, mas também aos insetos, uma maior biodiversidade.

O referido artigo de Otto Gotlieb (1992) traz a influência de conceitos da "Taxonomia Numérica" de Sokal (1973), assim definida no livro "*Systematics, Evolution, and Biogeography of Compositae*", de Funk (2009):

""*Taxonomia numérica mede o grau de alteração no estado de oxidação de "micromoléculas" (metabólitos secundários), indicando uma tendência no processo evolucionário que pode ser expressado como índices numéricos*""

Em relação à quimiotaxonomia, este livro traz a ideia geral de que a especialização que ocorre em estruturas químicas por meio de um crescente grau de oxidação, bem como a simplificação das cadeias carbônicas, acompanha a evolução das plantas. Além disso, sabe-se que muitas evoluções estão inter-correlacionadas entre diversos seres vivos, como plantas e insetos, e, deste modo, a quimiotaxonomia poderia ser explicada do ponto de vista prático, visto que ela pode evidenciar o relacionamento causa-consequência entre os metabólitos e o meio ambiente.

Mais precisamente em relação à caracterização química do gênero *Viguiera*, a grande maioria das substâncias isoladas tem sido as lactonas sesquiterpênicas (LST) da classe dos heliangolidos e seus derivados furanoeliangolidos e epoxieliangolidos, típicos do gênero (Da Costa, 2008). Demais caracterizações químicas de várias espécies do gênero foram relatadas, observando-se a ocorrência de LST dos tipos guaianolido e germacrolido (Schorr *et al.*, 2002), diterpenos, como os ácidos *ent*-caurenóico e *ent*-pimaradienóico, bem como os seus derivados (Da Costa *et al.*, 1996, Ambrosio *et al.*, 2004), e ainda flavonoides, poucas cumarinas e um novo derivado do mioinositol, fato único no gênero (Spring *et al*., 2001). Por outro lado, é rara a ocorrência de triterpenos, sendo que há poucos estudos com óleos essenciais e seus mono- e sesquiterpenos. Portanto, de tal maneira, a caracterização química das plantas é útil para o auxílio da taxonomia e filogenia, tal como pode ser observado em trabalhos realizados anteriormente com o gênero *Helianthus*, onde os autores propuseram inclusive alterações taxonômicas e de parentesco (Spring *et al*., 1989, 1990, 1991).

Além das contribuições quimiotaxonômicas, vale mencionar que muitos dos metabólitos secundários descritos acima são bioativos e, portanto, factíveis de serem estudados do ponto

de vista terapêutico (Verpoort, 1998; Ngo, 2013) e toxicológico (Schimidt, 199). Mais precisamente em relação ao potencial farmacológico do gênero *Aldama*, há diversos estudos que demonstram a importância deste gênero. Como exemplo, há um estudo muito interessante que avalia a atividade tripanomicida de diterpenos derivados do ácido *ent*-caurenóico (Da Costa *et al.,* 1996). Outro trabalho revelou a ação de diterpenos derivados da classe dos *ent*-pimaranos na ação inibidora da contratilidade da carótida de ratos (Ambrósio *et al.,* 2002; Tirapelli *et al.,* 2004). Outros trabalhos demonstraram a atividade antimicrobial do óleo essencial de *V. dentata* (Caneles *et al.,* 2008), bem como a atividade anticancerígena em ratos, desempenhada por LST isoladas de *A. sylvatica* (Taylor *et al.,* 2008).

Por fim, há artigos que demonstraram a importância dos metabólitos secundários em processos anti-inflamatórios (Gobbo-Neto *et al.,* 2005; Schorr *et al.,* 2005), alguns específicos ao gênero *Aldama* (Schorr *et al.,* 2002; Siedle *et al.,* 2004; Nicolete *et al.,* 2009). Recentemente também têm sido muito explorados os compostos fenólicos derivados dos ácidos cafeoilquínico e clorogênico, no que diz respeito à ação anti-inflamatória (Chagas-Paula *et al.,* 2011).

Com relação ao estudo das propriedades anti-inflamatórias de *Aldama* é importante demonstrar alguns dos processos bioquímicos que desencadeiam a inflamação. Sabe-se que há duas enzimas muito importantes envolvidas nos processos inflamatórios: as ciclooxigenases (COXs) e as lipoxigenases (LOXs), responsáveis por oxidar o ácido araquidônico, resultando na produção de prostanóides (prostaglandinas, prostaciclinas e tromboxanos) na via COX-dependente, ou leucotrienos, na via LOX-dependente (Rang e Dale, 2004). As prostaglandinas contribuem para os sintomas da inflamação aguda, tais como aumento da permeabilidade vascular, vasodilatação, formação de edema, produção de dor e febre. Já os leucotrienos são potentes quimiotáxicos e agentes broncoconstritores, e desta forma estimulam a infiltração de leucócitos, contribuindo para a ulceração por oclusão de micro vasos, bem como provocando crise de asma e conversão da inflamação aguda para crônica (Range e Dale, 2004; Parente, 2001). Os anti-inflamatórios não esteroidais (AINES) atuais são capazes de inibir as COXs e não são capazes de inibir as LOXs (Parente, 2001; Rang e Dale, 2004). Assim, a via de metabolização do ácido araquidônico é deslocada para ser metabolizado apenas pelas LOXs, resultando maior formação de leucotrienos, aumentando os efeitos adversos dos AINES (Rang e Dale, 2004; Parente, 2001; Fiorucci *et al*., 2001). Então, recentemente destacou-se a importância de procurar substâncias que sejam capazes de inibir tanto as COXs quanto as LOXs (Parente, 2001; Fiorucci *et al*., 2001; Charlier, 2003).

Deste modo, cabe aqui explicar o conceito de metabolômica e como ela pode nos auxiliar tanto na obtenção de dados químicos para a quimiotaxonomia como para realizar estudos de bioatividade. Metabolômica, por definição, seria a detecção de todos os metabólitos de um dado organismo, assim como genômica é o estudo do genoma completo de um dado organismo, e proteômica é o estudo de todas as proteínas e suas isoformas. O primeiro conceito de metaboloma introduzido (Oliver *et al*., 1998) menciona a necessidade de se fazer conexão entre metabolômica e genômica, observando-se as alterações de concentração do metabólito como um resultado da deleção ou expressão de algum gene. Logo em seguida (Fiehn *et al*., 2001 e 2002) a metabolômica é descrita como sendo "uma compreensiva e quantitativa análise de todos os metabólitos", sendo que tal abordagem revelaria o metaboloma de um sistema biológico e nos daria uma maior compreensão da genômica funcional. Já Goodacre e seus colaboradores (Goodacre *et al*., 2004) pela primeira vez listaram diferentes abordagens de metabolômica, embora concordem que ela deva ser quantitativa e complementar à transcriptômica e genômica. As diferentes abordagens definidas foram: *metabolite target analysis* (análise direcionada ao metabólito), *metabolite profiling* (perfil metabólito), *metabolomics* (metabolômica), *metabolic fingerprint* (impressão digital metabólica), *metabolic profiling* (perfil químico metabólico) e *metabonomics* (metabonômica). Porém, muitos destes conceitos já se alteraram desde 2004 e outros caíram em desuso. Por exemplo, a metabonômica, a qual fora colocada como sendo uma medida das alterações nos níveis dos metabólitos causados por perturbações causadas por drogas e toxinas, não tem sido mais recomendada. Metabolômica é colocada como uma ciência e não mais como uma abordagem, visto que ainda é impraticável a detecção e quantificação de todos os metabólitos. Assim sendo, os conceitos mais modernos do que seria metabolômica as descrevem como uma nova área da ciência, em vez de uma abordagem analítica (Villas-Bôas *et al*., 2005). Definem ainda que metabolômica é a caracterização do fenótipo metabólico (metaboloma) sob específicas condições (estágios de desenvolvimento do organismo, condições ambientais e genéticas, etc.) e a correlação ao seu correspondente genótipo (Villas-Bôas *et al*., 2005 e 2006; Summer *et al*., 2003; Schwab *et al*., 2003). Este fenótipo metabólico pode ser caracterizado sob as diferentes abordagens, onde a técnica do *metabolite profiling* (perfil metabólico) pode ser direcionada ou não. Nesta abordagem, objetiva-se a detecção dos metabólitos por um conjunto de técnicas analíticas específicas como, por exemplo, técnicas hifenadas baseadas em cromatografia e espectroscopia (Villas-Bôas *et al*., 2005 e 2006).

A metabolômica não-direcionada (*untargeted metabolomics*), que seria a mesma coisa que perfil metabólico não-direcionado, é uma técnica que visa à obtenção dos dados químicos de todo e qualquer metabólito, sem especificidade, por exemplo primário ou secundário. Metabolômica direcionada é quando, já de antemão, se propõe detectar determinada classe específica de um metabólito, ou uma substância química (ou derivados) em específico. Assim sendo, a extração fitoquímica de todo e qualquer metabólito pode ser considerada uma metabolômica não-direcionada, pois, embora o objetivo seja a detecção de metabólitos secundários, não se direciona a técnica para uma determinada classe química, como flavonoides, por exemplo.

Por fim, ainda no que diz respeito à metabolômica, é necessário descrever aqui conceitos que trazem uma relação direta entre as diferentes abordagens da metabolômica e a identificação e quantificação dos metabólitos, que por sua vez são inversamente proporcionais ao número de metabólitos detectados (Krastanov, 2010). Assim sendo, quanto mais conhecimento se tem sobre os tipos de metabólitos detectados, inclusive realizando a sua quantificação, mais a abordagem se aproxima de uma análise direcionada, ao passo que, quanto menos isso ocorre, mais a análise passa a ser classificada como de perfil metabólico ou até mesmo de impressão digital metabólica (Figura 3), a qual diz respeito apenas à aquisição dos espectros e cromatogramas dos perfis químicos. Estes conceitos não entram em conflito com o que fora exposto por Villas-Bôas, apenas descrevem melhor em termos quantitativos e qualitativos as abordagens metabolômicas.

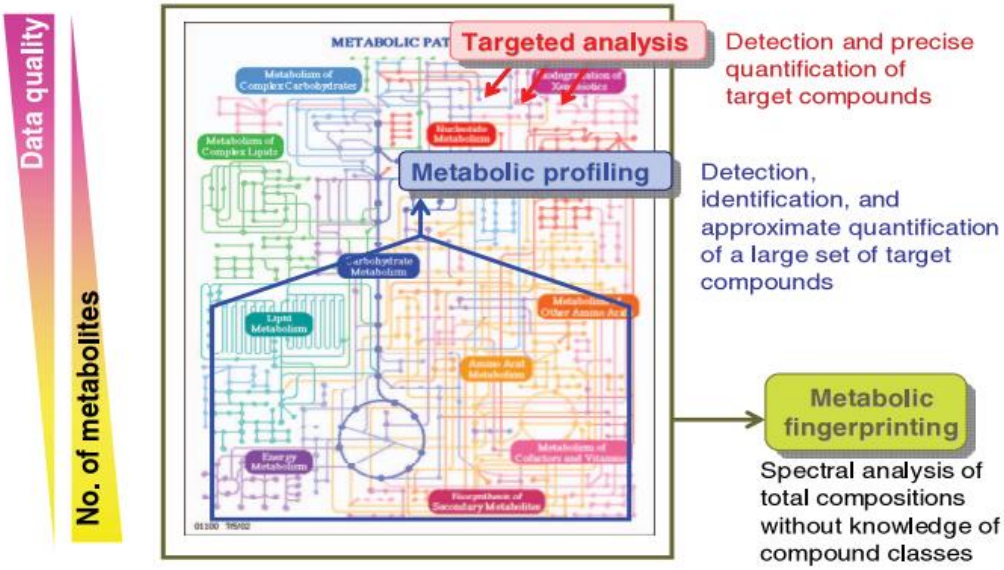


Figure 3 - Tipos de análises metabolômicas (Retirado de Krastanov, 2010; pág 1538).

Entretanto, devido à grande quantidade de dados químicos gerados pela abordagem da metabolômica não-direcionada, torna-se demasiado complexa e demorada a análise dos mesmos. Para tanto, é necessário a utilização de análise estatística multivariada supervisionada e não supervisionada, as quais estão inseridas na quimiometria, ramo da ciência que trata da aplicação de métodos estatísticos e matemáticos aos dados químicos, extraindo-se informações químicas de tais dados (Gasteiger e Engel, 2003). Os dados químicos são usualmente multidimensionais e, neste caso, são chamados de "multivariados". Os diferentes métodos matemáticos e estatísticos aplicados aos dados químicos podem ser a Análise das Componentes Principais (*Principal Component Analysis* – PCA), Análise de Agrupamento Hierárquico (*Hierarchical Cluster Analysis* – HCA), Mínimos Quadrados Parciais *(Partial Least Squares* – PLS) e sua variante ortogonal (OPLS), sendo que estas duas últimas podem ser utilizadas para Análises Discriminantes (*Discriminant Analysis* – DA), compondo as conhecidas técnicas PLS-DA e OPLS-DA. Assim sendo, a HCA e PCA são análises multivariadas não supervisionadas, pois elas são úteis para o reconhecimento de padrões, tendências e agrupamentos não direcionados das amostras. Já o PLS-DA é considerado uma análise multivariada supervisionada, pois o agrupamento observado segue uma tendência (um direcionamento) previamente informada à análise, por exemplo, as amostras se agrupam de acordo com uma específica informação fornecida à análise, a qual pode ser a bioatividade ou *clusters* formados por gêneros ou origem biogeográfica.

Melhor explicando, tanto a PCA quanto o PLS-DA são técnicas de redução de dimensionalidade (Rencher, 2002), onde um conjunto de amostras com variáveis multidimensionais são reduzidas no que se chama de Componentes Principais, no caso da PCA, e Variáveis Latentes, no caso do PLS-DA. Ambas técnicas projetam em um hiperplano bidimensional o que seria impossível de se observar em um espaço multidimensional. A PCA sempre traça a primeira componente na direção de maior variância encontrada entre as variáveis do sistema, ou seja, é justamente nesta direção que apenas uma componente principal teria maior precisão em explicar tudo aquilo que somente seria possível se observássemos o espaço multidimensional. Entretanto, embora a primeira componente sempre explique a maior variância encontrada, sempre é necessário mais de uma para explicar um espaço multidimensional. Assim sendo, uma segunda componente, ortogonal à segunda, explica uma maior quantidade de variância, menor que a primeira componente, mas maior que a terceira componente, à qual é ortogonal às duas primeiras. Quando maior a complexidade de um conjunto de dados multidimensionais, tão mais difícil é de se explicar com apenas duas

componentes a maior variância encontrada, sendo, portanto, necessários mais componentes. O PLS funciona da mesma forma, e o termo Componentes Principais é utilizado às vezes para designar as Variáveis Latentes. Entretanto, a diferença é que as Variáveis Latentes sempre apontam na direção de maior variância sobre uma determinada informação (a resposta em Y) fornecida ao sistema. Assim sendo, o PLS-DA correlaciona a variância do espaço multidimensional (a matriz X) com uma resposta informada (matriz Y) (Wold *et al*., 2001). Já o OPLS-DA é considerado "uma modificação do PLS, onde é realizada uma separação das variações da matriz X em duas partes, uma linearmente relacionada à matriz Y (componentes preditivas) e outra que é ortogonal à Y (não preditiva com a resposta Y) (Trygg *et al*., 2002; Verron *et al*., 2004; Manual SIMCA 13.0.3). Isso gera uma maior interpretabilidade, visto que as variações ortogonais a Y nos informam as diferenças dentro de X, mas que não se correlacionam com Y. Todas estas técnicas de redução dimensional do espaço multidimensional geram um novo gráfico (projeção do hiperplano bidimensional) de amostras, chamado agora de *scores*, os quais se relacionam diretamente com o novo gráfico gerado para as variáveis, chamado agora de *loadings*.

Já a HCA é um método de "*Cluster Analysis*" propriamente dito. Em outras palavras, a HCA necessariamente agrupa por similaridade as amostras do estudo, procurando por padrões e tendências dentro dos agrupamentos (*clusters*). Ela é representada na forma de uma "árvore" onde os *clusters* se formam ao longo de uma distância. Quanto mais próximo a 0, mais similar. Uma distância comum bastante utilizada é a Euclidiana, a qual trata-se da distância entre dois vetores, um x (x1, x2, ... , xp) e outro y (y1, y2, ... , yp), definido como :

$$d(x,y) = \sqrt{x} - y'(x-y) = \sqrt{\sum_{j=1}^{p}(xj - yj)(xj - yj)}$$

Há muitas formas de se fazer uma HCA, sendo as mais utilizadas o *Single Linkage*, o qual se utiliza da distância entre as amostras mais próximas para construir os *clusters*, e o método de Ward (*Ward's Method*), o qual usa a variância total dentro das amostras para definir os *clusters*. O resultado é que o método de Ward minimiza a somatória da distância dentro dos *clusters* (Rencher, 2002; Larose, 2006). A escolha de qual método se deve utilizar é empírica e depende das necessidades ou do que se pretende informar.

A quimiometria não se ocupa apenas da análise dos dados que visa gerar informação, mas também do processamento e tratamento dos mesmos, os quais são etapas anteriores às análises e cruciais na geração de adequados dados químicos. Por exemplo, os dados provenientes de cromatogramas e espectros geralmente necessitam de suavização da linha de base, alinhamento, normalização, entre outras técnicas de tratamento, visando a adequação dos dados químicos à posterior análise e geração de informação. Desta forma, para os dados químicos provenientes do LC-MS há diversos *softwares* de processamento de dados, dentre os quais daremos destaque ao *MZmine* (Katajamma *et al*., 2006; Pluskal *el al*., 2010) e *MetAlign* (Lommen, 2009; Lommen e Kools, 2012), entre outros, sendo que cada qual traz sua vantagem e desvantagem, como apresentado na Figura 4 (Castillo *et al*., 2011).

Data preprocessing steps covered by the proteomic and metabolomic software tools discussed in this paper.

Software	Filtering	Peak detection	Deisotoping	Alignment	Gap filling	Visualization	Identification	Normalization	Batch mode
apLCMS	✓	✓	–	✓	✓	–	–	–	✓
Decon2LS	✓	✓	✓	–	–	✓	–	–	–
Madison Metabolomics Consortium Database	–	–	–	–	–	–	✓	–	–
MAVEN	✓	✓	–	✓	–	✓	✓	–	–
MetAlign	✓	✓	–	✓	–	–	–	–	✓
MetaboAnalyst	–	✓	–	✓	✓	–	–	✓	–
MSight	–	–	–	–	–	✓	–	–	–
MZedDB	–	–	–	–	–	–	✓	–	–
MZmine 2	✓	✓	✓	✓	✓	✓	✓	✓	✓
OpenMS/TOPP	✓	✓	✓	✓	–	✓	✓	–	✓
Sirius	–	–	✓	–	–	✓	✓	–	–
SuperHirn	–	✓	✓	✓	✓	–	–	✓	–
XCMS	✓	✓	–	✓	✓	✓	✓	–	–

Figure 4 - Comparação de diferentes softwares de processamento de dados (Extraído de Castillo *et al*., 2011).

Os *softwares MZmine* e *MetAlign* são gratuitos, sendo que este último pode ainda ser utilizado em conjunto com o *software MSClust* (Tikunov *et al*., 2012), o qual traz algoritmos capazes de reduzir a quantidade de sinais de um espectrômetro de massas, oriundos de fragmentos moleculares, adutos e isótopos. Assim sendo, este *software* traz a promessa de agrupar os sinais em supostos metabólitos íntegros, podendo-se fazer identificação *a posteriori*.

Por fim, é necessário introduzir aqui os conceitos de *Data Mining* (Mineração de Dados) e de como ela tem sido utilizada pela quimiometria recentemente (Buydens et al., 1999; Mutihac, L. e Mutihac R., 2008), visto que a *Decision Tree*, uma de suas técnicas, é utilizada neste trabalho. Basicamente, os significados de *Data Mining* não diferem muito dos de quimiometria. *Data Mining* trata da resolução de problemas presentes em bancos de dados (os quais podem ser químicos), explorando-os por meio de técnicas computacionais, visando o reconhecimento de padrões e/ou tendências e classificações. Tais padrões nos permitem

realizar predições em novos dados. Atualmente, muitas das técnicas desenvolvidas em *Data Mining* podem ser utilizadas pela quimiometria, tais como Análise Bayesiana (*NaiveBayes*), Redes Neurais Artificiais e *Decision Trees* (Árvores de Decisão). *Decision Tree* é uma representação gráfica proveniente de "*divide-and-conquer*" (dividir e conquistar), um algoritmo que trabalha subdividindo problemas maiores em problemas menores que possam ser resolvidos (Quinlan, 1986 e 1987). Em uma *Decision Tree*, os "nós" da árvore representam um determinado atributo (uma variável) a ser testado por condições específicas (concentração, por exemplo). Tais testes fornecem as "folhas" da árvore, as quais contêm a classificação do teste (Witten, 2011). Há vários algoritmos que calculam uma *Decision Tree* e não convém explicá-los aqui, embora os mais utilizados sejam o C4.5, desenvolvido por Ross Quinlan (Quinlan, 1993), e o J.48 (Bhargava *et al*., 2013), implementação do C4.5 no *software* gratuito Weka (Hall, *et al*., 2009).

Por fim, pode-se destacar aqui trabalhos (Alvarenga *et al*., 2001; Emerenciano *et al*., 2006) que se utilizam da quimiometria para explicar a quimiotaxonomia de Asteraceae, descrevendo muito apropriadamente os metabólitos mais específicos de cada subtribo e gênero, sendo que o segundo trabalho traz aplicações diretas sobre a teoria redox de Gottlieb, mencionada anteriormente. Entretanto, estes trabalhos foram realizados com dados químicos provenientes da literatura e, portanto, se limitam apenas aos metabólitos isolados e identificados pela fitoquímica clássica. Aqui, nesta Dissertação, menciona-se a aplicação de quimiometria aos dados químicos provenientes de uma abordagem metabolômica não-direcionada.

Portanto, a quimiometria aplicada à metabolômica não-direcionada auxilia nos estudos quimiotaxonômicos do gênero, fundamentando-se no fato de que algumas classes de substâncias são específicas de alguns táxons (Sharma, 2007), possibilitando-nos fazer a desreplicação de substâncias discriminantes de determinados gêneros ou grupos infragenéricos. Além disso, esta combinação de técnicas nos permite direcionar a fitoquímica para metabólitos que seriam potencialmente ativos frente a um determinado alvo molecular ou enzimático, sendo que neste trabalho os alvos referem-se às enzimas anti-inflamatórias COX-1 e 5-LOX.

Objetivos

O objetivo deste trabalho foi realizar um estudo metabolômico não-direcionado (*untargeted metabolomics*) auxiliado por técnicas quimiométricas (*chemometrics-assisted*), objetivando à desreplicação de metabólitos secundários específicos de *Aldama* La Llave e/ou *clusters* infragenéricos, contribuindo para a quimiotaxonomia do gênero. Além disso, tal estudo visa também à indicação específica de substâncias químicas potencialmente ativas contra as enzimas COX-1 e 5-LOX, possibilitando uma fitoquímica direcionada (*targeted phytochemistry*).

Para melhor entendimento, os objetivos podem ser subdivididos nos seguintes tópicos:

- obtenção dos dados químicos das espécies de *Aldama* La Llave por meio da abordagem da metabolômica não-direcionada (*untargeted metabolomics*), ou inespecífica;
- análises multivariadas visando à criação de modelos quimiométricos específicos para estudos em quimiotaxonomia e inibição concomitante das enzimas COX-1 e 5-LOX;
- desreplicação direcionada (*targeted dereplication*) para a quimiotaxonomia e de substâncias potencialmente ativas contra COX-1 e 5-LOX;
- fitoquímica direcionada (*targeted phytochemistry*) para isolamento e elucidação estrutural de substâncias potencialmente ativas contra as enzimas COX-1 e 5-LOX, possibilitando ulterior ensaio *in vitro*, visando confirmar tal bioatividade.

Material e Métodos

1 – Coleta e Secagem

Foram coletadas – ou cedidas por integrantes de um Projeto Temático (processo FAPESP nº 2010/51454-3) ao qual esta Dissertação está inserida – as seguintes espécies: *Aldama anchusifolia (DC.)* E.E.Schill. & Panero, *A. arenaria (Baker)* E.E.Schill. & Panero, *A. aspillioides* (Baker) E.E.Schill. & Panero, *A. bakeriana* (S.F. Blake) E.E.Schill. & Panero, *A. bracteata* (Gardner) E.E.Schill. & Panero, *A. discolor* (Blaker) E.E.Schill. & Panero, *A. filifolia* (Sch.Bip. ex Baker*) E.E.Schill. & Panero*, *A. gardneri* (Baker) E.E.Schill. & Panero, *A. goyazii* E.E.Schill. & Panero, *A. grandiflora* (Gardner) E.E.Schill. & Panero, *A. kunthiana* (Gardner) E.E.Schill. & Panero, *A. linearifolia* (Chodat) E.E.Schill. & Panero, *A. megapotamica* (Malme) Magenta & Pirani, *A. nudibasilaris* (S.F.Blake) E.E.Schill. & Panero, *A. nudicaulis* (Baker) E.E.Schill. & Panero, *A. oblongifolia* (Gardner) E.E.Schill. & Panero, *A. pilosa* (Baker) E.E.Schill. & Panero, *A. robusta* (Gardner) E.E.Schill. & Panero, *A. rubra* (Magenta & Pirani) E.E.Schill. & Panero, *A. santacatarinensis* (H.Rob. & A.J.Moore) E.E.Schill. & Panero, *A. squalida* (S.Moore) E.E.Schill. & Panero, *A. tenuifolia* (Gardner) E.E.Schill. & Panero, *A. trichophylla* (Dúsen) Magenta, *A. tuberosa* (Griseb.) E.E.Schill. & Panero. Além destas espécies, também foram adquiridas duas variedades (ou espécimes de outras populações) de *A. robusta* e *A. tuberosa*, sendo a *A. robusta* var. *oxylepis* (S.F.Blake) E.E.Schill. & Panero e *A. tuberosa* var. *guaranitica* (Chodat.) E.E.Schill. & Panero. Os locais de coleta de cada espécie de *Aldama* estão descritos no Apêndice 5 (Locais de coleta de *Aldama* La Llave).

Figure 5 - População de *A. trichophylla* (foto de minha autoria)

Foram estudadas apenas as folhas e tricomas glandulares das folhas de todas as plantas adquiridas ou coletadas somente em épocas de florada (período de máxima maturidade da planta) e no período da manhã (anterior às 12h00). As folhas são o órgão vegetal onde classes de metabólitos de maior interesse se localizam, como terpenos e compostos fenólicos; além disso, também é possível dar prosseguimento aos estudos fitoquímicos, onde uma maior quantidade de material é indispensável. Para se preparar os extratos, as plantas foram secadas em estufa de ar circulante, a 40 ºC, durante 24 h, no Laboratório de Farmacognosia da FCFRP-USP. Também é importante mencionar que este projeto possui autorização do Ministério do Meio Ambiente para a realização de coletas de material vegetal (nº 36266-1).

2 – Extrato Total (ET)

O Extrato Total consiste no extrato de folhas íntegras, previamente secadas e pulverizadas em moinho analítico de facas. O ET foi obtido pela extração de 20 mg da droga vegetal em 2 mL de EtOH-H_2O 70% por maceração em agitação orbital em *shaker* orbital (CT-713RT – CIENTEC) por 24 h, 25 °C e 120 rpm. Após isso, foram agitados em vórtex e centrifugados em microcentrífuga (CT-15000 – CIENTEC) a 14.000 rpm por 3 min. Os sobrenadantes dos extratos foram coletados e secados em *SpeedVac Concentrator* (Savant SPD2010 – *Thermo Electron Corporation*) e, posteriormente, ressuspendidos em 400 µL de MeCN-H_2O 1:1 (v/v). Para se remover o material graxo e pigmentos não relevantes ao estudo, os extratos foram particionados com *n*-hexano em uma relação de 1:1 (v/v). Os Extratos Totais livres de material graxo foram filtrados em filtros PTFE de 20 µm (Chromafil O-20/15 MS – *Macherey-Nagel GmbH & Co. KG*) anteriormente à injeção no LC-DAD-MS (*Liquid Chromatography – with Diodo-Array Detector – coupled to Mass Spectrometer*). Este projeto possui autorização do CNPq (nº 010055/2012-6) para acesso ao patrimônio genético.

3 – Coleta de Tricomas Glandulares (TG)

Foram coletados 100 TG presentes nas folhas e dissolvidos em 100 µL de MeCN:H_2O (1:1 v/v). Foram armazenados em freezer -20 °C até o momento da injeção no UHPLC-DAD-Orbitrap (*Ultra High Pressure Liquid Chromatography – with Diodo-Array Detector – coupled to Orbitrap*™ *Mass Spectrometer*).

As espécies das quais se coletaram os TG são apenas aquelas onde estes foram encontrados em quantidade suficiente para a análise, sendo elas: *A. rubra*, *A. discolor*, *A. robusta*, *A. robusta* var. *oxylepis*, *A. pilosa*, *A. anchusifolia*, *A. tenuifolia*, *A. kunthiana*, *A.*

bakeriana, *A. gardneri*, *A. arenaria*, *A. arenaria* (inflorescências – TGI). Nas Figuras 6.a e 6.b podem ser observadas fotografias dos TG de folhas e nas Figuras 7.a e 7.b os tricomas glandulares presentes nas anteras das inflorescências de *A. arenaria*.

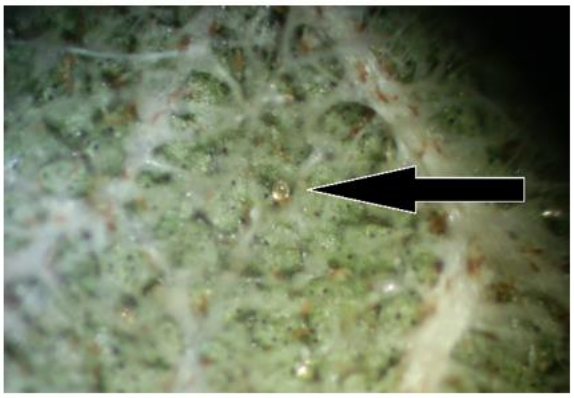


Figura 6.a – Tricoma Glandular em folha de *A. robusta*

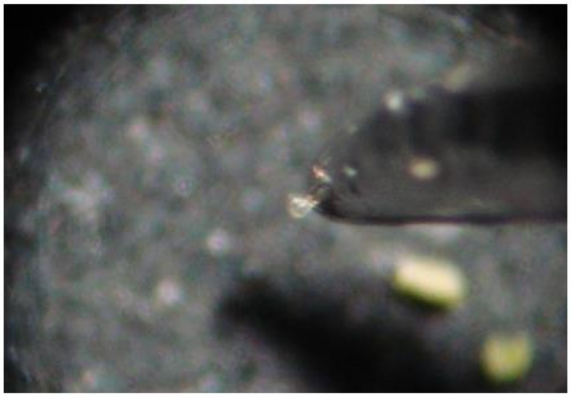


Figura 6.b – Tricoma Glandular pinçada por agulha

Fotos de minha autoria

Figura 7.a – Anteras de inflorescência de *A. arenaria*

Figura 7.b – Tricomas Glandulares em anteras de *A. arenaria*

Fotos de minha autoria

4 – Extrato de Lavagem Foliar (ELF)

Foi obtido o Extrato de Lavagem da superfície das folhas secas e íntegras por meio da lavagem de 2,5 g de folhas em 150 ml de EtOH por aproximadamente 30 seg, a fim de se extrair todo o conteúdo dos tricomas glandulares, presentes na superfície das mesmas, porém com o mínimo possível de material graxo ou pigmentos, não inerentes a este estudo. O solvente foi rotaevaporado e o extrato seco foi ressuspendido em MeOH. Em seguida foi realizada partição com *n*-hexano, visando-se a remoção de material graxo para a injeção no UHPLC-DAD-Orbitrap. Vale mencionar que apenas as espécies do item anterior de “Material e Métodos” é que

passaram pela lavagem foliar. Também foi realizada lavagem em inflorescências de *A. arenaria* (ELI) do mesmo modo que fora realizado o ELF.

5 – Obtenção dos dados químicos via HPLC-PDA-TOF e UHPLC-DAD-Orbitrap^{TM} e processamento dos dados via MZmine, MetAlign e MSClust.

Os ET, TG e ELF foram analisados via UHPLC-DAD-Orbitrap para as contribuições quimiotaxonômicas. Já nos estudos de inibição enzimática foi utilizado apenas o ET, visto que este foi o extrato utilizado nos ensaios *in-vitro*, o qual foi analisado via HPLC-DAD-TOF (*High Liquid Chromatography – with Diodo-Array Detector – coupled to Time-of-flight Mass Spectrometer*) e UHPLC-DAD-Orbitrap.

Para as análises em HPLC-DAD (*Shimadzu*) foram utilizadas duas colunas monolíticas de fase reversa (Onix C-18, 3 x 100 mm, 130 Å de poro, *Phenomenex*) acopladas entre si e um analisador de massas TOF (micrOTOF – II – *Bruker Daltonics*). O gradiente de eluição utilizado foi: bomba A: H_2O (1 % de ácido acético); bomba B: MeCN; 0 min: 0% B, 60 min: 45 % B, 65 min: 100%, 70 min: 100% B, 72 min: 0% B, 77 min: 0% B; a vazão da faze móvel foi de 1,4 mL/min.

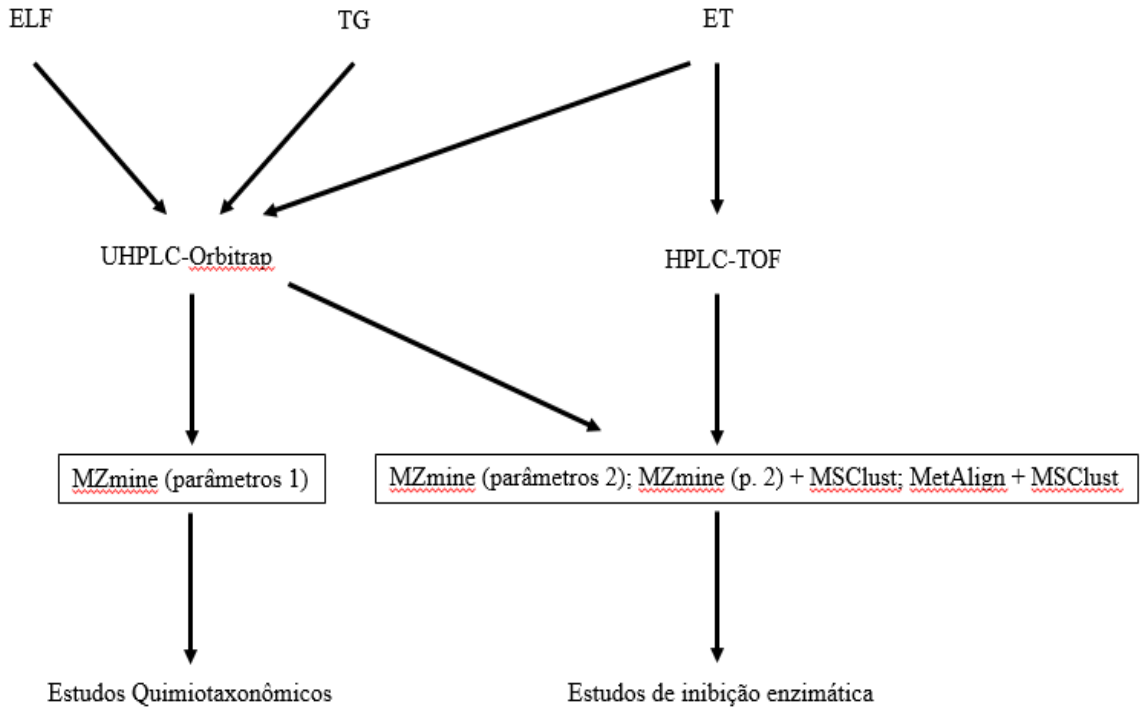
Para as análises em UHPLC-DAD (Accela) foi utilizada uma coluna cromatográfica de fase reversa (C-18, 150 x 3 mm, 3 µm de diâmetro de partícula ACE®), protegida por uma pré-coluna equivalente e analisador de massas Orbitrap^{TM} (Exactive Plus – *Thermo Scientific*). A vazão da fase móvel foi de 0,4 mL/min, utilizando H_2O com ácido fórmico (0,01%) (v/v) na bomba A e MeCN na bomba B, de acordo com o seguinte gradiente linear de eluição: 0 – 17 min, 2 - 55% B; 17 – 20 min, 55 - 100% B; 20 – 24 min, 100% B; 24 – 27 min, 100 - 2% B; 27 – 30 min, 2% B. Utilizou-se temperatura controlada em 35 °C para a cromatografia realizada no sistema Accela.

Em ambos os espectrômetros de massas os dados foram adquiridos nos modos de ionização positivo e negativo por *electrospray*, com 3,5 kV de energia de ionização (diferença de potencial de um capilar e um contraeletrodo). A resolução do TOF-MS foi da ordem de 16.500, enquanto que a resolução do Orbitrap-MS foi selecionada em 70.000.

Para os dados provenientes do UHPLC-Orbitrap, os parâmetros e algoritmos processados no *MZmine 2.10* foram utilizados segundo o APÊNDICE 1 (Parâmetros 1 do *MZmine 2.10*), os quais haviam sido previamente selecionados pela Mestre Danniela Príscylla Vasconcelos Faleiro, membro do referido Projeto Temático, e seguiu-se para as análises

quimiotaxonômicas. Para os dados provenientes do HPLC-TOF, os quais foram utilizados exclusivamente nas análises de inibição enzimática, os parâmetros e algoritmos processados no *MZmine 2.9* foram utilizados segundo o APÊNDICE 2 (Parâmetros 2 do *MZmine 2.9*), os quais foram comparados com os parâmetros do *MetAlign* (versão 041012) otimizados pelo *MSClust*, bem como pelos parâmetros do *MZmine* otimizados pelo *MSClust*. Assim, fez-se três comparações aqui: *MZmine vs MetAlign + MSClust vs MZmine + MSClust*. Os dados provenientes do UHPLC-Orbitrap também foram analisados nestas condições somente no estudo da inibição enzimática, permitindo-se comparar com o HPLC-TOF. Para que isso fosse possível, os dados provenientes do Orbitrap, gerados em arquivos *.raw*, tiveram que ser transformados em arquivos *.mzXML*, para que pudessem ser novamente rodados no *MZmine* utilizando-se os Parâmetros 2 (APÊNDICE 2), os quais são os mesmos em que os dados do HPLC-TOF foram processados. Além disso, diferentes parâmetros do *MZmine* foram testados em todas as diferentes análises em que ele foi utilizado. Estes diferentes parâmetros do *MZmine 2.9* escolhidos para comparação foram: *Filtering* (*Savitzky-Golay vs Mean Filter*), *Mass Detection* (*Centroid vs Wavelet Transform*) e *Deconvolution* (*Baseline Cut-off vs Local Minimum Search*). O modo negativo do TOF apresentou muito ruído, portanto optou-se pela escolha do *Noise Amplitude* e não do *Local Minimum Search*. O parâmetro FTMS (*Fourier Transform Mass Spectrometry*) *Lorentzian Extended* foi utilizado apenas para o Orbitrap. Estas combinações realizadas no *MZmine 2.9* podem ser melhor entendidas no Esquema 2, o qual mostra as oito combinações formadas para ambos analisadores de massas e os respectivos códigos de cada uma no modo de ionização positivo, embora tenham sido realizados para o modo negativo também.

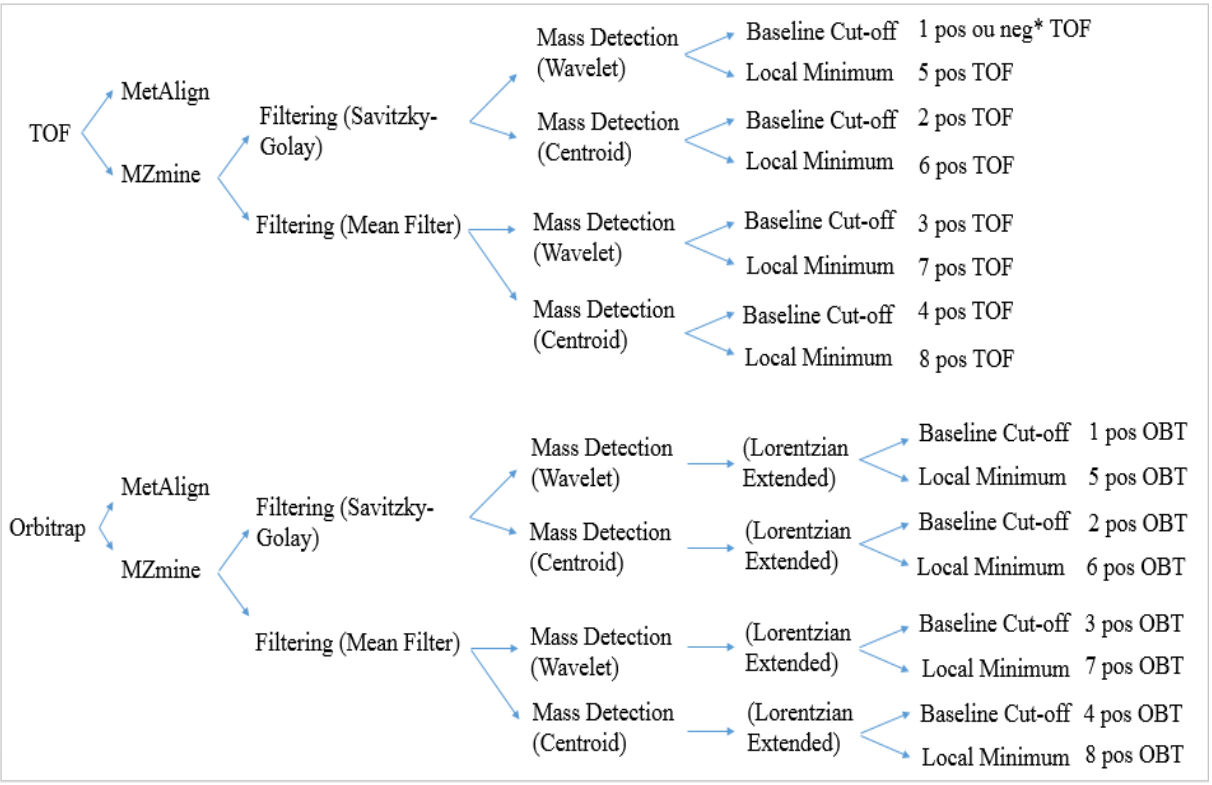
Todas estas combinações de diferentes técnicas hifenadas, parâmetros/algoritmos e *softwares* diferentes foram realizadas visando-se escolher aquela que fornecesse o melhor modelo de predição, capaz de informar com maior grau de certeza as substâncias responsáveis pela dupla inibição de COX-1 e 5-LOX. Todas estas combinações podem ser melhor entendidas observando-se o Esquema 1.



Esquema 1 – Esquema geral da obtenção e processamento dos dados químicos dos extratos vegetais de *Aldama* em software de processamento de dados de LC-MS oriundos do TOF e do Orbitrap.

O *MetAlign* (versão 041012) e o *MSClust* foram rodados segundo os parâmetros e algoritmos dos Apêndices 3 (Parâmetros do *MetAlign* (versão 041012)) e 4 (Parâmetros do *MSClust*), respectivamente.

Também é importante mencionar aqui que, embora muitos dos parâmetros dos softwares descritos acima sejam escolhidos de forma empírica, eles ainda devem estar em consonância com o tipo de analisador de massas empregado (poder de resolução, resolução, acurácia, etc.), bem como o perfil cromatográfico ou o TIC (*Total Ion Chromatogram*) adquiridos, e de acordo com a literatura (Holcapek, *et al*., 2012).



Esquema 2 - Diferentes algoritmos utilizados no software MZmine 2.9. pos: ionização no modo positivo; neg*: indica que o mesmo esquema foi utilizado para os dados provenientes para do modo negativo.

6 – Análise in vitro *do potencial anti-inflamatório dos extratos totais (ET)*

Para a análise do potencial anti-inflamatório foram utilizados *kits* de *screening* de inibidores da COX-1 (nº de catálogo 560101, *Cayman Chemical Company*, EUA) e 5-LOX (nº de catálogo 760700 e 60401, *Cayman Chemical Company*, EUA). A triagem dos extratos inibidores foi realizada de acordo com as instruções dos fabricantes dos *kits*. Na ocasião, este experimento foi realizado em conjunto com a Dra. Daniela Aparecida Chagas de Paula, M.Sc. Danniela Príscylla Vasconcelos Faleiro (membros do referido Projeto Temático da FAPESP) e a doutoranda Rosana Casoti, todos membros da equipe científica AsterBioChem (www.asterbiochem.org) do Laboratório de Farmacognosia da FCFRP-USP.

A produção de prostaglandina (PG) foi dosada em imunoensaio (ELISA), baseado na competição entre PG livre (produzidas) e PG conjugada à enzima colinesterase (PG-conj. – que faz parte do *kit*) pelo limitado número de anticorpos anti-PG incorporado na placa de 96 poços. As placas de 96 poços foram lavadas para remover reagentes não ligados aos poços pelos anticorpos anti-PG incorporados aos poços das placas. O reagente de Elman, que contém substrato para acetilcolinesterase, foi adicionado a cada poço. O produto da reação

enzimática da acetilcolinesterase é amarelo e a intensidade desta cor é lida em leitor de placas no comprimento de onda de 405 nm. A intensidade da cor é proporcional à quantidade de PG-conj ligada; desta forma, uma maior densidade ótica (DO) indica maior poder de inibição da produção de PGs (PGs livres) por inibição da COX pelos extratos. A quantidade de PGs foi mensurada por interpolação da DO obtida em curva de DO por concentração de PGs. As porcentagens de inibição foram determinadas de acordo com a seguinte fórmula:

$$\text{Porcentagem de inibição} = \frac{PGc - PGt}{PGc} \times 100$$

Onde PGc é a quantidade de PGs do controle sem tratamento e PGt é a quantidade de PGs do extrato testado. Após isso foi realizado o cálculo de IC_{50}.

Para avaliar a inibição da via LOX-dependente foi seguido um protocolo de acordo com as recomendações do *kit*. A atividade da LOX é mensurada pela produção de hidroperóxidos que absorvem no comprimento de 490 nm, onde menores densidades óticas indicam maior inibição das lipoxigenases (5-LOX). As porcentagens de inibição foram determinadas de acordo com a seguinte fórmula:

$$\text{Porcentagem de inibição} = \frac{DOc - DOt}{DOc} \times 100$$

onde DOc é a densidade ótica do controle sem tratamento e DOt é a densidade ótica do extrato testado. Após isso foi realizado o cálculo de IC_{50}. O controle positivo para COX-1 foi indometacina e para 5-LOX foi o ácido nordihidroguaiarético (NDGA), com concentrações de atividade de 0,6 e 0,8 µg/mL, respectivamente.

É importante mencionar que a análise do potencial de inibição das enzimas COX-1 e 5-LOX foi realizada com extratos de apenas 17 espécies de *Aldama*, descritas na Tabela 2 (item 1 de “Resultados e Discussão”).

7 – Quimiometria dos dados químicos provenientes de LC-DAD-MS

Antes de serem inseridos nos *softwares* de quimiometria, os dados químicos utilizados foram normalizados em 1, quando havia presença de determinado metabólito, e 0, quando havia ausência. Os dados químicos utilizados nos modelos de inibição enzimática foram normalizados pela amplitude, segundo a fórmula:

$$A = \frac{x - xmin}{xmáx - xmin}$$

onde x é o valor da área sob o pico de um determinado metabólito, ionizado e detectado, a ser normalizada, e x_{min} e $x_{máx}$ são os valores mínimo e máximo da área sob o pico de um determinado metabólito, respectivamente.

Para a realização da PCA e OPLS-DA foi utilizado o *software* SIMCA 13.0.3 (Umetrics – Suécia); já a HCA foi obtida com o *The Unscrambler X 10.3* (CAMO™, Noruega). A *Decision Tree* foi realizada com Weka 3.6 *Data Mining Software* (The University of Waikato – Nova Zelândia). A PCA e HCA foram utilizadas para uma análise prévia dos dados, verificando-se a presença de *outliers*, padrões e tendências de agrupamento das amostras.

É importante mencionar que no software SIMCA os dados químicos, já normalizados como descrito anteriormente, foram também ajustadas pela técnica de Pareto (*Pareto scaling*), os quais revelam-se muito úteis em estudos metabolômicos (Manual SIMCA 13.0.3). O cálculo deste ajuste é:

$$Par = \frac{x - \bar{x}}{\sqrt{SD}}$$

Novamente, x são as áreas sob o pico de cada metabólito e SD é o desvio padrão. A consequência da utilização de *Pareto scaling* é que valores muito baixos (ou iguais a 0) necessariamente se tornarão negativos. Quanto mais baixo um determinado valor, se comparado com outros, mais negativo. Assim sendo, este cálculo melhora as visualizações nas análises multidimensionais, pois os *scores* e *loadings* ficarão distribuídos por todos os quatro quadrantes do gráfico, e não apenas nos quadrantes positivos do eixo das abscissas do hiperplano bidimensional. Por outro lado, no que diz respeito à amplitude dos valores, este ajuste tem pouca influência. Na Figura 8 do manual do SIMCA 13.0.3 pode-se observar a transformação que o *Pareto scaling* causa em um dado bruto de espectrometria de massas.

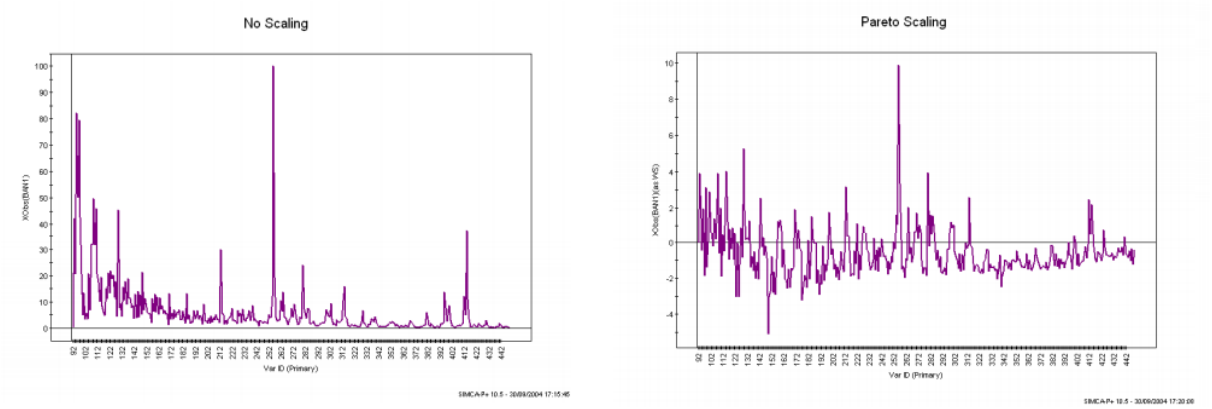


Figura 8 - Exemplo do *Pareto scaling* no tratamento de dados químicos provenientes de EM.

O OPLS-DA foi utilizado com a finalidade de apontar as substâncias químicas correlacionadas com uma determinada resposta informada à análise multivariada supervisionada. Por exemplo, para os estudos quimiotaxonômicos foi informada a classificação infragenérica de Blake (1918), a qual utilizou-se de parâmetros morfoanatômicos, classificação biogeográfica de acordo com climas específicos de determinadas regiões do Brasil (Alvares *et al*., 2014) dos locais de coletas de cada espécie de *Aldama*, e confronto com a quimiodiversidade frente a outros gêneros, os quais estão descritos no APÊNDICE 6 (Outros Gêneros). Para o estudo da inibição dupla de COX-1 e 5-LOX *in-vitro* foi utilizada a informação das espécies que apresentaram a atividade de inibição dupla e aquelas que não apresentaram tal atividade.

8 – Desreplicação direcionada (Targeted dereplication)

Por desreplicação direcionada (*Targeted dereplication*) aqui mencionada, pretende-se informar que a desreplicação foi realizada apenas para os picos detectados (prováveis substâncias químicas) pelo EM, os quais se apresentaram como discriminantes por meio das análises via OPLS-DA.

Visando facilitar a desreplicação dos dados provenientes do HPLC-TOF, foi utilizado o HPLC-Ion-Trap-MS, o qual permite fragmentar íons específicos por meio de uma amplitude de radiofrequência específica. Por meio do analisador Ion-Trap (amaZon SL – *Bruker Corporation*) foi possível fazer um estudo de monitoramento de reações múltiplas (MRM) até o estágio de MS^3 (terceira geração de íons fragmento gerados a partir de um íon precursor

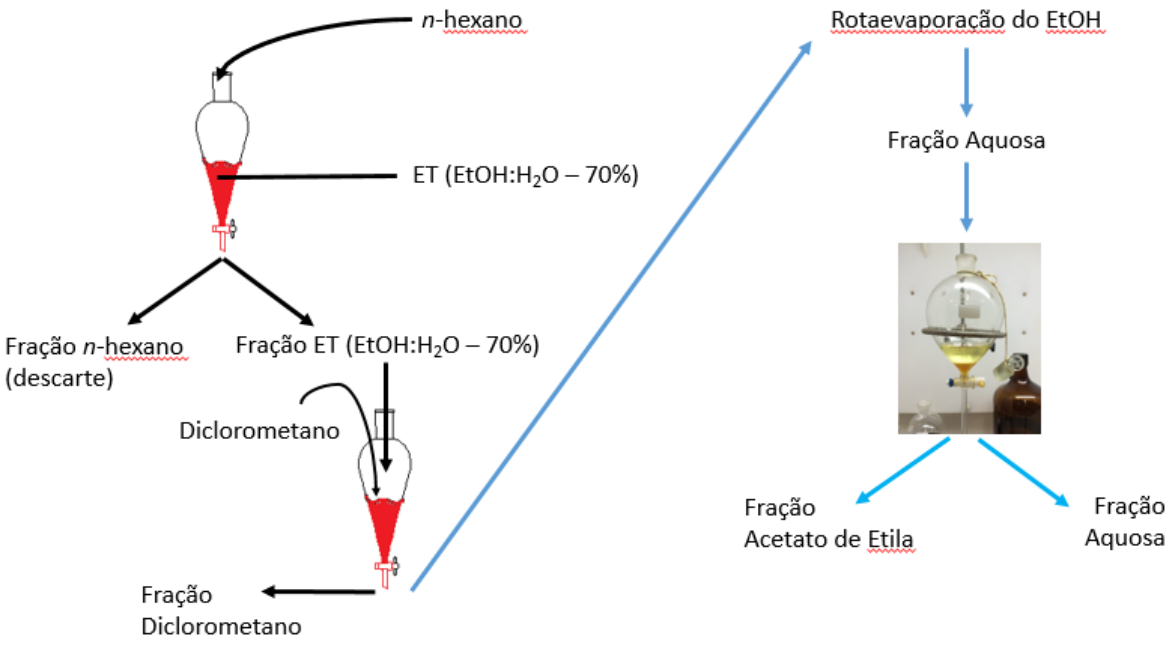
(*parent ion*)). A ionização do LC-Ion-Trap foi por *electrospray* (ESI), com energia de 3,5 kV. A amplitude de fragmentação foi de 0,6 e 0,5 para gerar os íons MS^2 e MS^3, respectivamente. Quando necessário, esta amplitude foi aumentada para 0,7 nos casos em que o íon não estava fragmentando. Já para as análises realizadas por meio do UHPLC-Orbitrap foi utilizada a cela *Higher-energy Collisional Dissociation* (HCD) com energia de colisão normalizada (*Normalized Collision Energy* – NCE) da ordem de 35% para se fazer a fragmentação dos íons, objetivando-se uma desreplicação mais ágil, uma vez que, na NCE a energia de fragmentação aumenta à medida que a relação *m*/*z* também aumenta, possibilitando uma fragmentação mais uniforme e nenhuma perda por *low-mass cut-off*.

Dados de UV da literatura e bancos de dados *in-house* de tempo de retenção, bem como um banco de massas protonadas e desprotonadas, adutos e dímeros de 313 substâncias químicas isoladas de *Aldama* La Llave, pesquisadas no *Scifinder Scholar*®, também foram utilizados na desreplicação.

9 – Isolamento dos picos alvos (Targeted phytochemistry)

Os picos alvos, apontados pelas análises quimiométricas como sendo aqueles mais provavelmente relevantes, do ponto de vista estatístico, na inibição dupla das enzimas COX-1 e 5-LOX, foram escolhidos para se fazer o isolamento direcionado. Além destes picos, outros de alta concentração também foram escolhidos para o isolamento via métodos de fitoquímica clássica. Tal como explicado no item 5 de "Resultados e Discussão", a planta de escolha para a fitoquímica foi a *A. trichophylla*, ainda muito pouco estudada do posto de vista fitoquímico.

Primeiramente, foi necessário fazer uma nova extração nas mesmas condições do item 2 de "Material e Métodos", porém, desta vez com quantidade de folhas suficiente para se realizar posterior isolamento e, portanto, em uma extração mais concentrada. Desta forma cerca de 250 g de folhas secas foram colocadas em maceração com 2 L de EtOH:H_2O (70% v/v) por 24 h. Após isso, este extrato (ET) foi particionado por solventes orgânicos de acordo com o Esquema 3.



Esquema 3 - Partições líquido-líquido do ET.

Para se utilizar um menor volume de solvente orgânico no particionamento do ET, sempre que necessário e possível, o EtOH e água foram concentrados em rotaevaporador R-210 (BÜCHI – Suíça). Os solventes orgânicos foram concentrados em rotaevaporador Fisatom (Brasil).

Após análise prévia em LC-DAD-MS verificou-se que apenas a fração AcOEt continha os picos apontados pela quimiometria (prováveis substâncias químicas de acordo com a desreplicação – ver item 4 de Resultados e Discussão). Assim sendo, 1,5 g da partição (fração) Acetato de Etila foram ressuspendidos em 3 mL de EtOH, filtrados em PVDF 0,45 µm (Analítica) e posteriormente submetidos ao fracionamento em uma coluna de vidro de 32 cm contendo *Sephadex* LH-20 (Figura 9.a) visando o isolamento das substâncias referentes aos picos. A fase móvel utilizada foi EtOH e coletou-se 70 frações de 15 mL por meio de um coletor automático (Pharmacia LKB – Frac-200 – Figura 9.b). Estas frações foram analisadas em CCD analítica (Whatman) com a seguinte fase móvel: AcOEt:MeOH:H_2O:HAc (80:5:5:15 v/v). O revelador químico escolhido foi solução ácida de anisaldeído (0,5 mL de anisaldeído, 10 mL de HAc glacial, 15 mL de MeOH e 5 mL de H_2SO_4 concentrado).

Figura 9.a – *Sephadex*

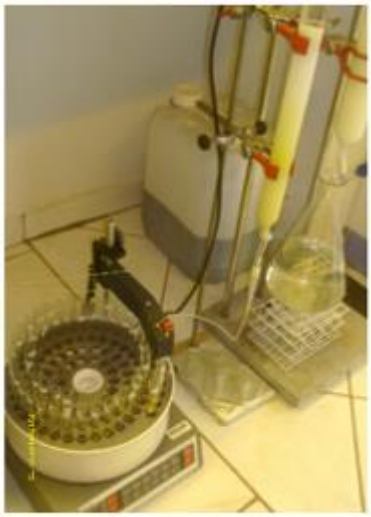


Figura 9.b – Coletor automático

Após a análise em CCD analítica, as frações foram reunidas em 12 novas frações de acordo com o perfil químico apresentado . Estas novas frações foram renomeadas como S-1 (fr. 1 e 2), S-3 (fr. 3), S-4 (fr. 4), S-5 (fr. 5), S-7 (fr. 6-11), S-12 (fr. 12-15), S-16 (fr. 16 e 17), S-19 (fr. 18-21), S-27 (22-31), S-37 (fr. 32-39), S-45 (fr. 40-70) e foram analisadas via HPLC-PDA-TOF e UHPLC-PDA-Orbitrap nas mesmas condições do item 5 de "Material e Métodos". Feito isso, as frações mais promissoras (item 5 – "Resultados e Discussão") foram submetidas a isolamento via HPLC-semi-preparativo (*Shimadzu*) utilizando-se uma coluna C-18 monolítica de 100 x 10 mm (130 Å de poro, Onyx – *Phenomenex*) e a seguinte fase móvel: 0 – 10 min, 0 – 20% B; 10 – 20 min, 20 – 25% B; 20 – 25 min, 25 – 45% B; 25 – 27 min, 45 – 65% B; 27 – 29 min, 65 – 65% B; 39 – 30 min, 65 – 0% B; 30 – 33 min, 0 – 0% B.

Os picos isolados foram novamente secados em *SpeedVac Concentrator* (Savant SPD2010 – *Thermo Electron Corporation*) e posteriormente seguidos para elucidação estrutural em RMN de acordo com o próximo item.

10 – Elucidação estrutural via RMN

Foi realizado o RMN ^{1}H em 500 MHz. Para tanto, todas as substâncias isoladas foram ressuspendidas em 500 µL de MeOD (*Cambridge Isotope Laboratories, Inc.*), independentemente da massa obtida no isolamento.

Resultados e Discussão dos estudos quimiotaxonômicos

1 – Hierarchical Clustering Analysis (HCA) *e Principal Component Analysis* (PCA)

Os estudos de quimiotaxonomia foram realizados com as espécies e gêneros de Asteraceae descritos no APÊNDICE 6 (Outros Gêneros) e com todas as espécies de *Aldama* La Llave descritas no item 1 de "Material e Métodos". É importante lembrar aqui que os dados químicos provenientes do UHPLC-Orbitrap, os quais são as áreas dos picos dos íons detectados, foram normalizados como "presença e ausência", sendo que, quando determinado íon era encontrado em uma dada espécie, este valor foi igualado a 1, significando presença, e 0, para o caso de ausência. Por hora, esta abordagem tem sido considerada pelo nosso grupo de pesquisa como a mais adequada, em vez de se trabalhar com dados de concentração, mesmo que normalizados pela amplitude, como descrito em "Material e Métodos". A abordagem da presença e ausência assume que, não importando em quais concentrações um determinado metabólito fora encontrado na planta, para ele ter sido biossintetizado, evidentemente, foi necessária uma maquinaria enzimática e, portanto, genes que o codificassem, podendo melhor serem correlacionados com a proteômica e genômica em casos de estudos pontuais. Em outras palavras, trata-se de uma análise metabolômica não-direcionada exclusivamente qualitativa, e não quantitativa, uma vez que, assim sendo, não há correlação da concentração dos metabólitos com qualquer tipo de variação ambiental, tais como estudos circadianos, sazonais, de interações ecoquímicas, etc., mas apenas a correlação com o genoma em uma determinada época pontual que, neste caso, foi a época de florada pelo período da manhã (até antes do 12h00). Esta preocupação em ser fiel à proteômica e genômica diz respeito ao fato de que, recentemente, árvores filogenéticas e taxonômicas têm sido construídas com base nestes estudos. Diante disso, poderíamos tentar comparar a Análise por Agrupamento Hieráquico (*Hierarchical Clustering Analysis* - HCA), construída com dados químicos e demonstrada nesta Dissertação, com tais classificações taxonômicas, mas sempre tendo em mente a limitação maior desta abordagem, uma vez que os dados químicos podem ser alterados por condições ambientais, o que fatalmente acarretariam em alterações nos *clusters* apresentados pela HCA.

Ainda com relação a HCA, uma outra dúvida que tem recentemente surgido nos grupos de pesquisas que trabalham com dados químicos adquiridos via espectroscopia de massas é se ela deve ser construída exclusivamente com dados provenientes do modo de ionização

positivo, negativo, ou ambos. Para alguns, talvez parece demasiado problemático em não se juntar os modos de ionização, uma vez que determinados metabólitos se ionizam apenas no modo positivo ou negativo e, assim sendo, a HCA poderia não representar fielmente uma classificação quimiotaxonômica. Para outros, a preocupação é inversa e diz respeito ao fato de alguns metabólitos se ionizarem em ambos os modos e, portanto, serem contados duas vezes para algumas espécies que os produzem, propiciando que elas ficassem em um *cluster* mais próximo. Infelizmente não há dados na literatura que encorajem explicar os modos de ionização separadamente e juntos em um mesmo manuscrito para que pudéssemos nos basear ou comparar. Desta forma, parece-nos mais apropriado mostrar aqui ambas abordagens, sem pré-conceitos.

Ao analisar as HCA realizadas com os modos de ionização separadamente (Figuras 11 e 12), percebe-se que aquela proveniente do modo de ionização negativo não conseguiu unir todas as espécies do gênero *Aldama* em um único *cluster*. Pelo modo de ionização positivo, foi observado que *A. bakeriana* e *A. aspillioides* ficam fora do *cluster* de *Aldama*, o qual está mais unido desta vez. Poderíamos, de fato, supor que do ponto de vista químico, o gênero se subdivida em dois grupos no modo de ionização negativo. Assim, assumiríamos que os metabólitos com grande tendência a perder próton, formando moléculas desprotonadas ($[M-H]^+$), tais como os flavonoides, formam dois grandes grupos dentro do gênero *Aldama*. Infelizmente, por hora esta é apenas uma hipótese, visto que ainda não foi possível realizar esta desreplicação. Quando à HCA do modo positivo, outra hipótese que poderíamos levantar ao tentar explicar uma maior união do gênero é o fato de, neste modo, as LST, além de se ionizarem se fragmentam, aumentando a aproximação das espécies de *Aldama*; a contra-hipótese a isso seria o fato de não encontrarmos com facilidade LST em todas as espécies.

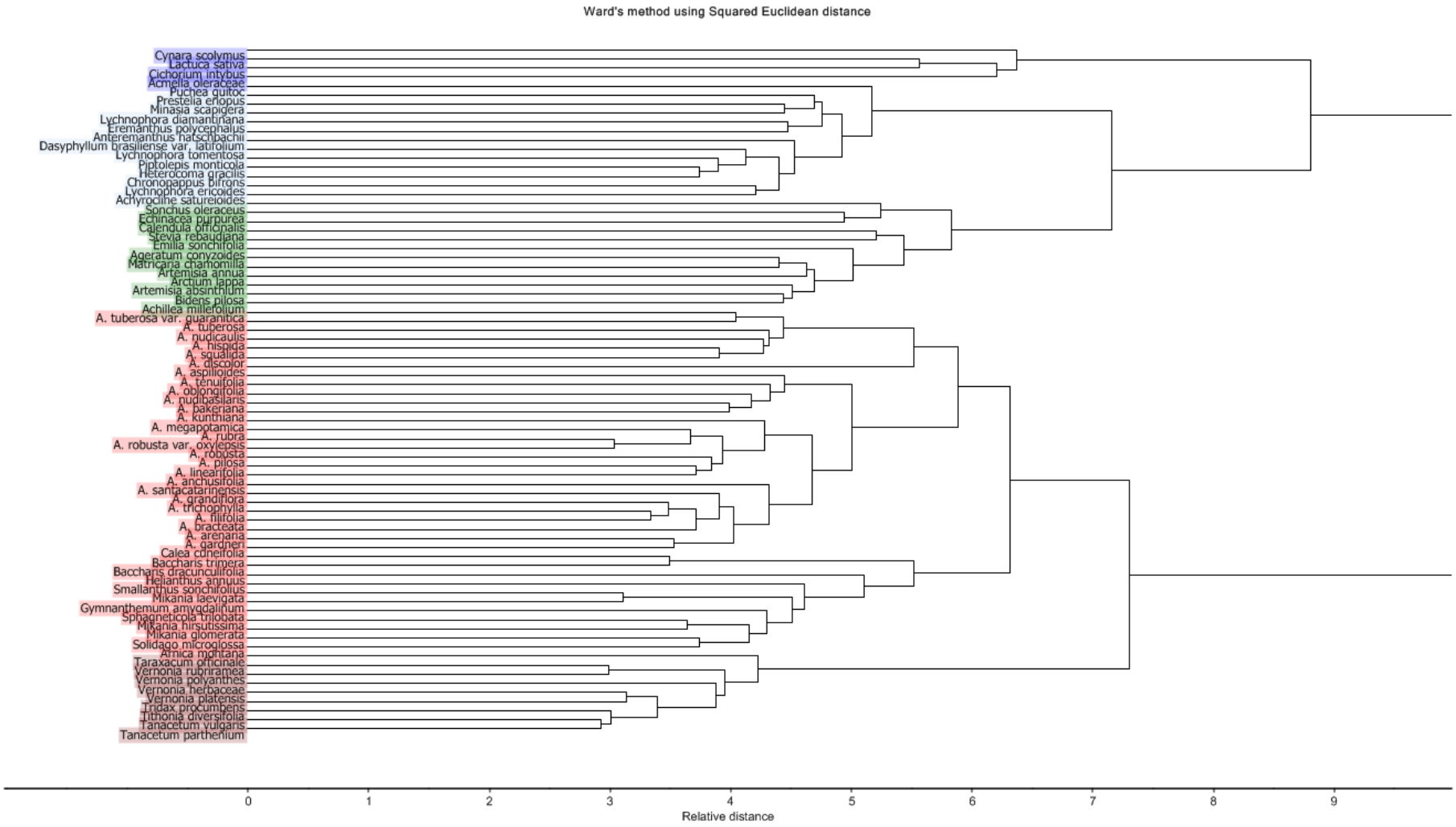


Figura 10 - HCA realizada com dados químicos provenientes dos modos de ionização positivo e negativo do ET de *Aldama* La Llave (em vermelho) e todos os outros gêneros de Asteraceae.

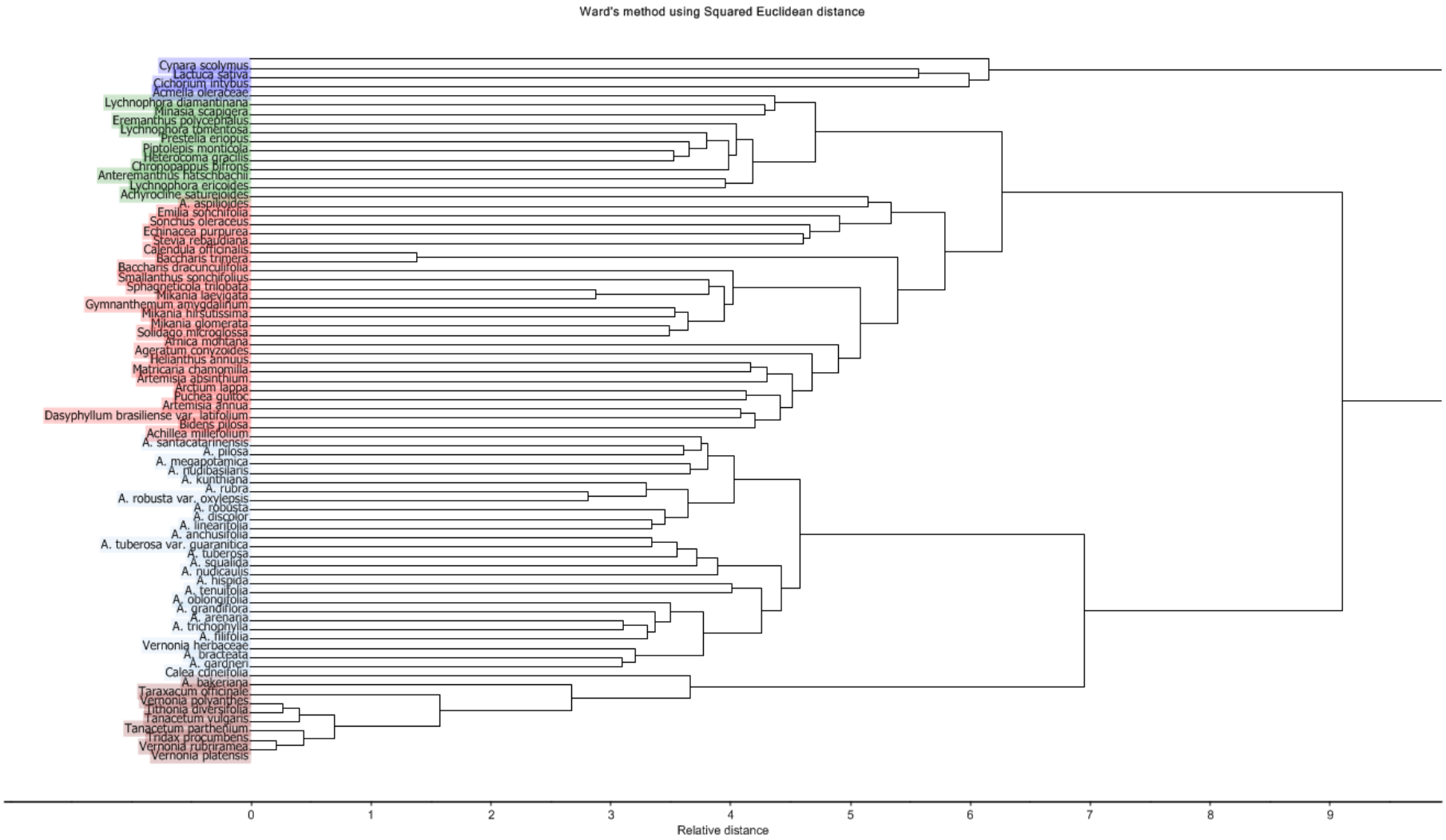


Figura 11 - HCA realizada com dados químicos provenientes do modo de ionização positivo do ET de *Aldama* La Llave (em azul) e todos os outros gêneros de Asteraceae.

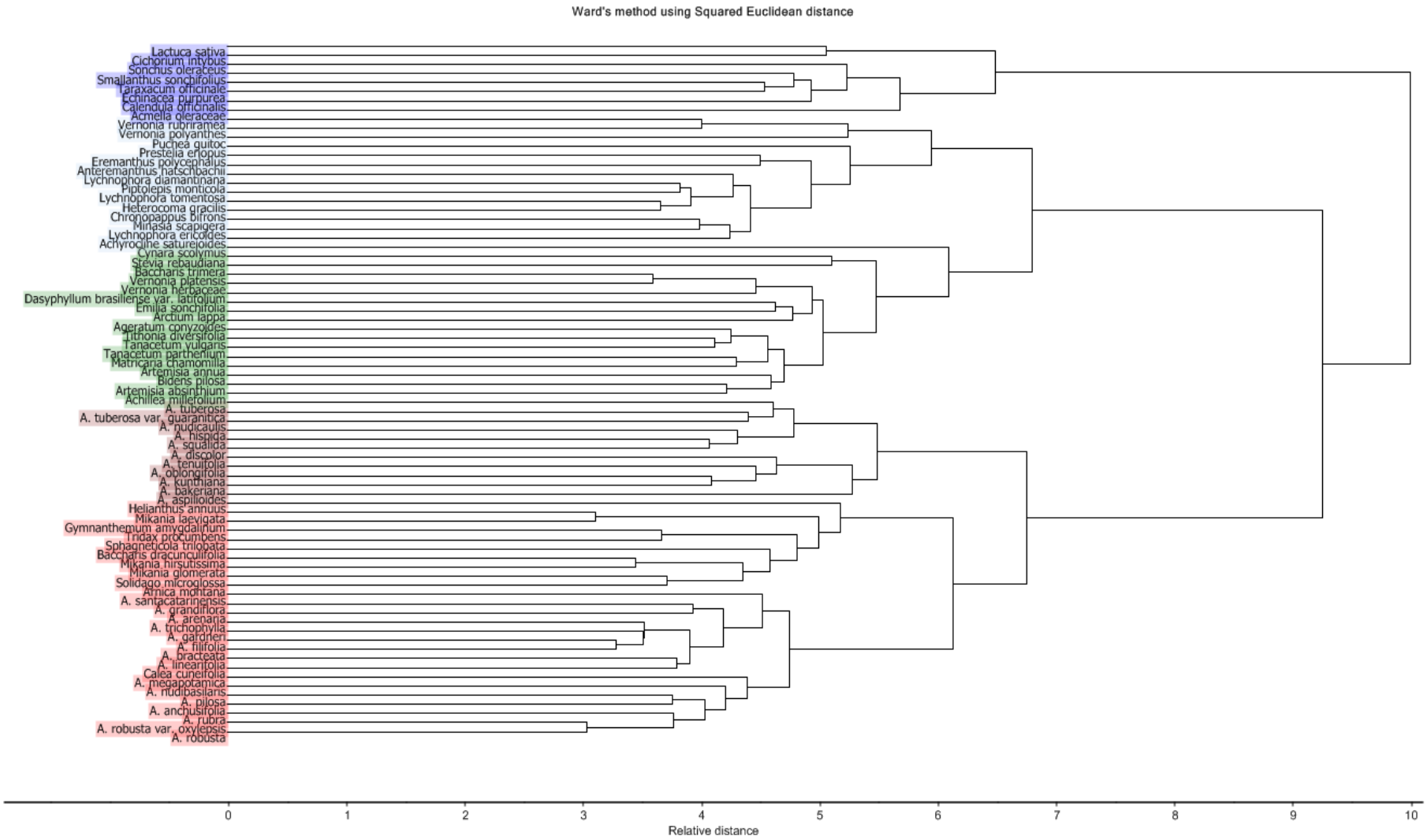


Figura 12 - HCA realizada com dados químicos provenientes do modo de ionização negativo do ET de *Aldama* La Llave e todos os outros gêneros de Asteraceae.

Por outro lado, a HCA realizada com ambos os modos de ionização juntos (Figura 10), mostra os agrupamentos entre todas as espécies de *Aldama versus* outros gêneros de Asteraceae. Por meio desta análise, pode-se perceber a quimioconsistência de gênero *Aldama*, sendo que todas as espécies permanecem unidas, em um *cluster* bem definido. É importante salientar aqui que não é o propósito da quimiotaxonomia estabelecer novos clados, agrupamentos, visto que, tal como mencionado, a constituição química das plantas pode sofrer influências ambientais que, neste trabalho não foram monitoradas. Além disso, há que considerar as eventuais limitações inerentes à técnica de UHPLC-DAD-ESI-Orbitrap. Em outras palavras, a detecção química é dependente da capacidade do metabólito em sofrer ionização via ESI sem, entretanto, se fragmentar ou formar de adutos, os quais poderiam comprometer as análises, visto que eles por si só já seriam considerados múltiplos de um único metabólito, um íon precursor. Outra limitação da técnica é a supressão iônica, que é quando determinados metabólitos não se ionizam devido ao efeito de matriz, por exemplo. Além disso, sabemos das dificuldades dos metabólitos apolares em se ionizarem por ESI (Zhou *et al*., 2012).

Ainda assim, a principal ideia da utilização da abordagem "quimiometria aplicada à metabolômica não-direcionada" é auxiliarmos em eventuais dificuldades encontradas pela taxonomia e filogenia. Por falar em filogenia, deve-se lembrar que esta agrupa os organismos com base no grau evolutivo genético, os quais podem ter correlação com as similaridades fenotípicas que, neste trabalho, tratam-se dos constituintes químicos. Outro intuito destas classificações químicas por meio de técnicas quimiométricas é apontar e evidenciar os "porquês" de tais agrupamentos observados. Eis aqui a vantagem da química sobre a genômica, ou mesmo a proteômica, uma vez que ela nos traz informações diretas das relações ambientais entre os seres vivos. Este conceito era muito bem compreendido por Gottlieb:

"*O maior desafio da ciência neste novo século é tentar entender os mecanismos de funcionamento da natureza. A Química tem um papel fundamental nessa história (...). Através dela é possível entender a origem, a diversidade e o futuro da vida*".

Tal frase consta como contracapa do periódico *Química Nova*, vol. 35, nº 11, 2012, e suas ideias podem ser bem compreendidas no artigo "Químico-biologia quantitativa: um novo paradigma?" (Gottlieb e Borin, 2012), publicado no mesmo periódico.

Para apontarmos os constituintes químicos responsáveis por determinados agrupamentos, nós podemos nos utilizar das análises multidimensionais, tais como PCA e

OPLS-DA, por meio das quais o espaço multidimensional de todas as variáveis seriam projetados em um hiperplano bidimensional, onde poderíamos correlacionar as amostras dos extratos vegetais (*scores*) com as variáveis químicas (*loadings*). A técnica de OPLS-DA será abordada no próximo item, e a PCA (Figura 13) foi utilizada aqui para fins classificatórios não supervisionados, ou seja, os *loadings* que poderiam se correlacionar com um determinado *score* não é específico dele (discriminante). Deste modo, não fizemos a desreplicação destes metabólitos na PCA, mas apenas por meio do OPLS-DA. Além disso, através da PCA nós podemos localizar *outliers* e escolher se os deletaríamos de nossas análises, antes de se fazer o OPLS-DA.

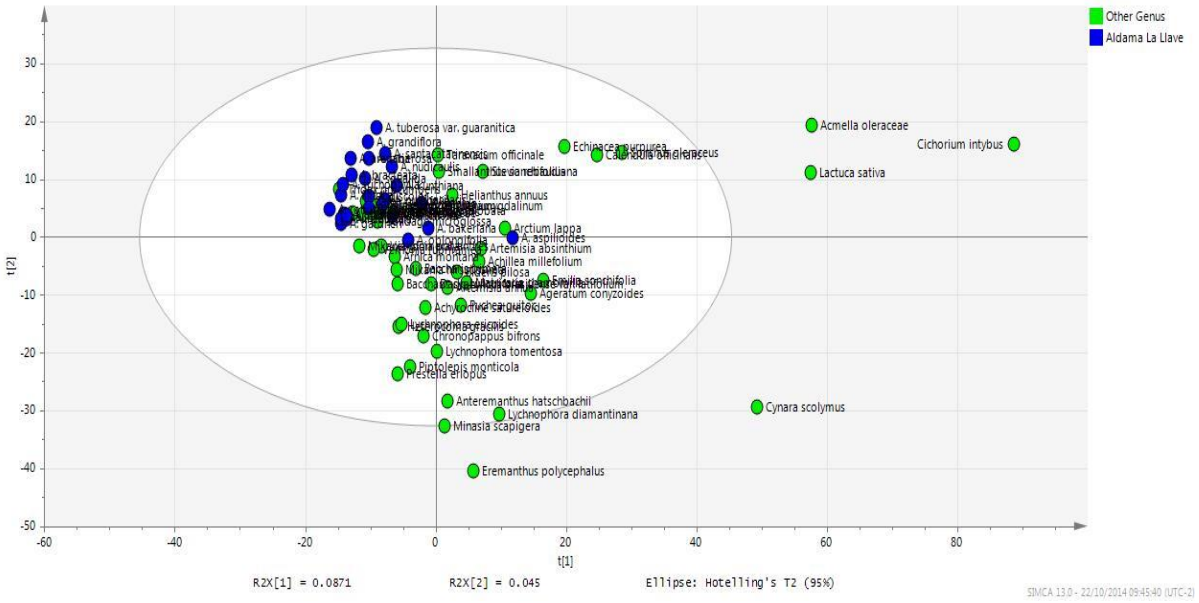


Figura 13 - PCA com dados provenientes de ambos os modos de ionização de ET de *Aldama* La Llave e outros gêneros

Deve ser informado que todas as análises multidimensionais foram realizadas com ambos os modos de ionização concomitantemente. Isso é permitido em análises multidimensionais, uma vez que, parafraseando um conceito físico, aqui duas variáveis podem ocupar um mesmo espaço em um gráfico de *loadings*, contribuindo, portanto, a um mesmo *score*. Em outras palavras, tais *loadings* não são contados duas vezes e eles possuem a mesma contribuição na classificação dos *scores*.

Na literatura há poucos trabalhos que descrevem a utilização de quimiometria aplicada à metabolômica e, mais especificamente em relação a esta abordagem em estudos quimiotaxonômicos. Recentemente Martucci e colaboradores (2014) concluem em seu

trabalho que a utilização de metabolômica não-direcionada pode ser aplicada em quimiotaxonomia, complementarmente às classificações morfoanatômicas. Este resultado é novo, pois não se utiliza apenas de dados químicos provenientes da fitoquímica clássica, os quais estão documentados na literatura, para realizar estudos quimiotaxonômicos, tais como aqueles demonstrados pelos trabalhos de Alvarenga e colaboradores (2001) e Emerenciano e colaboradores (2005), para tribos e gêneros de Asteraceae.

Aqui, além de demonstrarmos a quimioconsistência do gênero *Aldama* La Llave em relação a outros gêneros de Asteraceae, nós poderíamos tentar correlacionar os *clusters* específicos deste gênero por meio de uma comparação com a classificação taxonômica infragenérica proposta por Blake (1918), realizada com base em estudos morfoanatômicos. Blake subdivide o gênero em subgêneros, seções e séries, tal como pode-se observar na Figura 14, onde o gênero *Aldama* está com seu nome antigo, *Viguiera*.

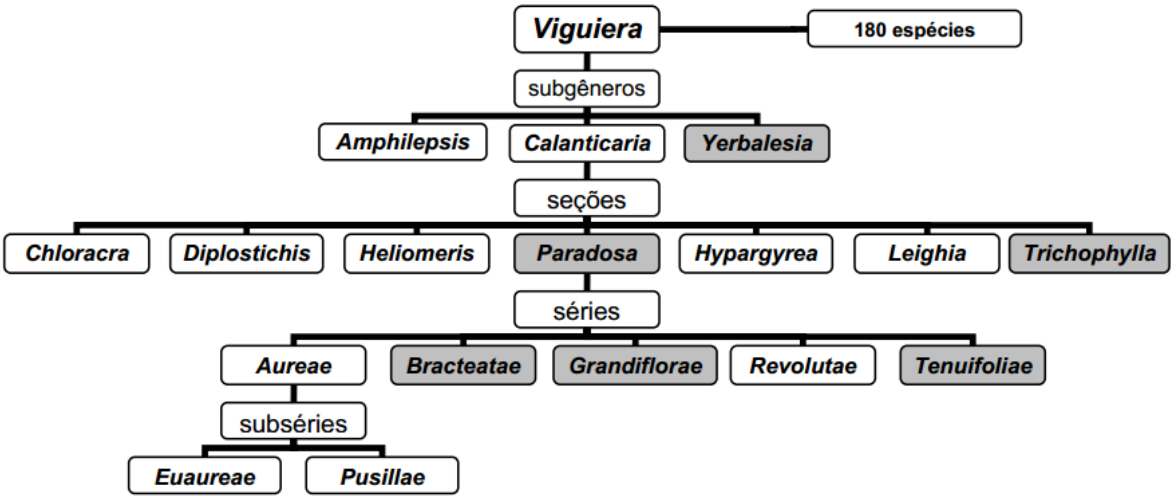


Figura 14 - Classificação de Blake (1918) (Extraído de Da Costa, 2008 – Texto Sistematizado de Livre Docência).

Aqui é feita a comparação entre a classificação quimiotaxonômica e as classificações de Blake porque, por hora, ainda não há um estudo que demonstre uma árvore filogenética das espécies brasileiras analisadas neste projeto. Tal estudo está sendo realizado por um projeto de pós-doutorado intitulado "Filogenia das espécies sul-americanas de *Viguiera* (Asteraceae: Heliantheae)", o qual faz parte do referido Projeto Temático da FAPESP. Assim sendo, na Figura 15 pode-se observar a HCA ampliada para as espécies de *Aldama*.

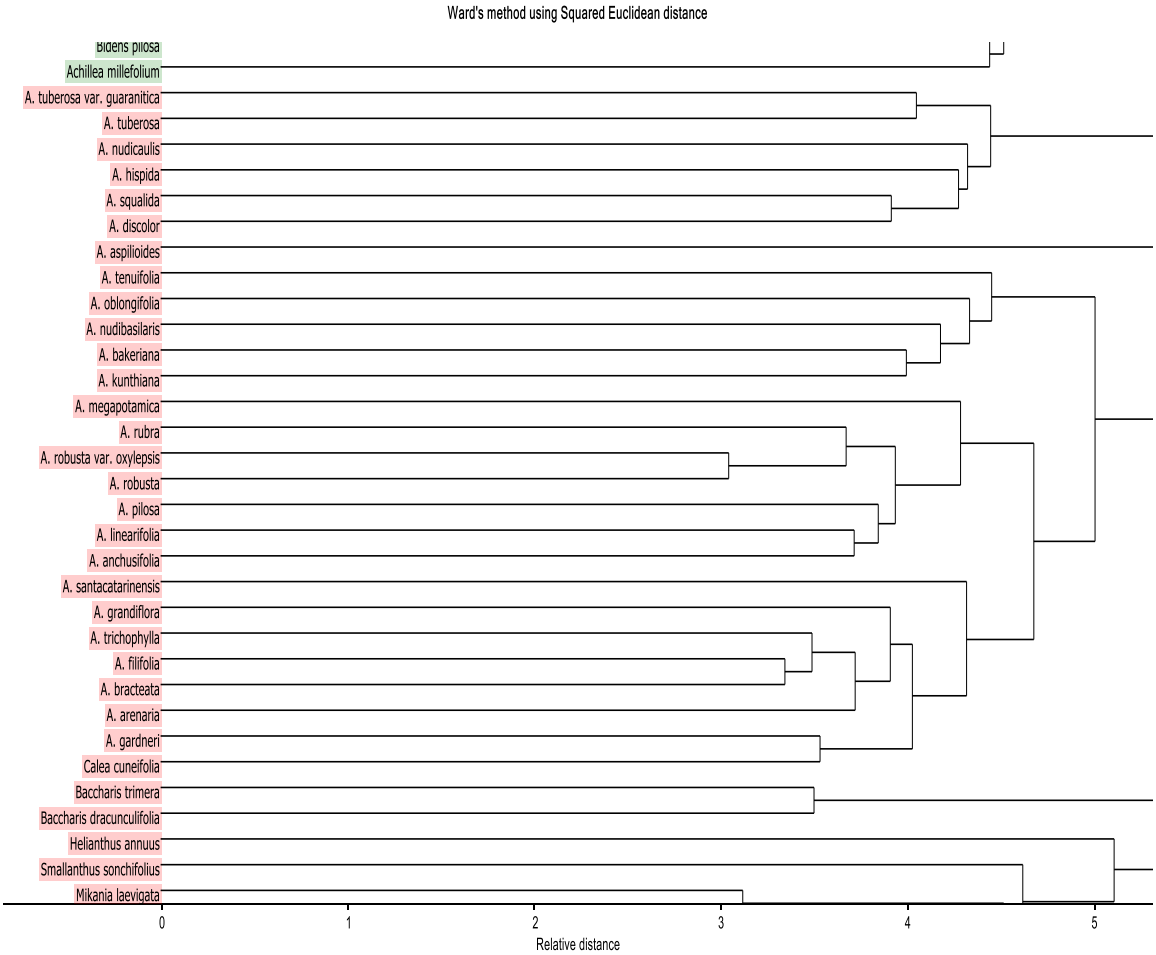


Figura 15 – HCA com base nos dois modos de ionização, ampliada, para o gênero *Aldama* La Llave (em vermelho).

Para poder comparar com os *clusters* da Figura 15 tem-se que primeiramente informar a quais seções e séries pertencem as respectivas espécies. Isso pode ser observado no Quadro 1. Pode-se perceber que três espécies não entram na classificação de Blake (*A. rubra*, *A. megapotâmica* e *A. santacatarinensis*). Todavia, tais espécies poderiam ser preditas em tais seções e séries com base em seus constituintes químicos e um modelo de experimental OPLS-DA. Isto será apresentado no próximo item. Por enquanto, atentemo-nos em relação à HCA da Figura 15, por meio da qual pode-se observar que há pouca correlação entre os dois estudos. Embora isso possa parecer um pouco frustrante, a química de uma planta não deve necessariamente se correlacionar com seus caracteres morfoanatômicos. De fato, ela poderia estar se correlacionando com qualquer outro fator externo a ela, como os ambientais. É evidente que os fenótipos morfoanatômicos também se correlacionam com o meio ambiente, e, assim como os fenótipos químicos, também se correlacionam com a proteômica e genômica, mas é um erro lógico assumir que, necessariamente, os caracteres

morfoanatômicos devam se correlacionar com os químicos. Além disso, não apenas devemos levar em consideração que tais estudos de Blake são muito antigos e, certamente, possuíam equívocos, mas também devemos lembrar que a nossa abordagem é a metabolômica não-direcionada, que leva em consideração todos os metabólitos detectados de acordo com a técnica analítica empregada. Assim sendo, no próximo item, há um estudo de OPLS-DA que, sendo uma técnica supervisionada, procura no espaço multidimensional o agrupamento relacionado com a classificação de Blake.

Quadro 1 - Classificação infragenérica de Blake (1918).

Lista das espécies classificadas por Blake (1918)	
A. anchusifolia	Aureae
A. hispida	Aureae
A. nudibasilaris	Aureae
A. pilosa	Aureae
A. arenaria	Bracteatae
A. bracteata	Bracteatae
A. gardneri	Bracteatae
A. oblongifolia	Bracteatae
A. robusta	Bracteatae
A. robusta var. *oxylepsis*	Bracteatae
A. bakeriana	Grandiflorae
A. discolor	Grandiflorae
A. grandiflora	Grandiflorae
A. squalida	Grandiflorae
A. rubra	Other_1
A. santacatarinensis	Other_2
A. megapotamica	Other_3
A. aspilioides	Tenuifolia
A. kunthiana	Tenuifolia
A. tenuifolia	Tenuifolia
A. filifolia	Trichophylla
A. linearifolia	Trichophylla
A. trichophylla	Trichophylla
A. nudicaulis	Yerbalesia
A. tuberosa	Yerbalesia
A. tuberosa var. *guaranitica*	Yerbalesia

Outro detalhe que podemos observar na HCA apresentada anteriormente é que as variedades (ou diferentes populações), as quais são grupos infraespecíficos e, portanto,

servem de parâmetro de controle de qualidade deste estudo (como se fosse padrão interno), ficaram próximas às suas respectivas espécies. Também havíamos informado na Introdução os problemas taxonômicos de *Aldama* La Llave, como a confusão que é feita em relação a *A. arenaria* e *A. robusta*. Aqui, tais espécies ficaram distantes, em *clusters* bem separados. Tal distinção também fora noticiada em um trabalho que descreve as diferenças nas estruturas secretórias de tais espécies (Tuane Oliveira *et al*., 2013). Além disso, *A. anchusifolia* e *A. megapotâmica* também se encontram em *clusters* bem separados aqui neste estudo. Quanto às espécies *A. linearifolia*, *A. filifolia* e *A. trichophylla*, pode-se perceber que a primeira se localiza em um *cluster* separado das duas últimas. De fato, parece haver muita semelhança química apresentada pelo agrupamento da *A. trichophylla*, *A. filifolia* e *A. bracteata*, o qual está na mesma distância Euclidiana (mesmo nível) apresentada pelas *A. robusta*. Ademais, fora introduzido nesta Dissertação que *A. discolor*, *A. squalida* e *A. bakeriana* puderiam se tratar da mesma espécie. Nossos resultados evidenciam que *A. bakeriana* se encontra em um cluster bem distinto das demais, ao passo que, de fato, *A. discolor* e *A. squalida* estão muito próximas do ponto de vista da constituição metabolômica.

2 – Orthogonal Partial Least Squares – Discriminant Analysis (OPLS-DA) e desreplicação dos metabólitos discriminantes.

Tendo em vista que as análises não supervisionadas apresentadas no item anterior, aqui podemos mostrar os resultados do modelo de predição de acordo com a classificação de Blake (1918). Para tanto, primeiramente foi feita uma PCA (Figura 16) com as mesmas espécies de *Aldama* do item anterior, por meio da qual observou-se que a espécie *A. aspillioides* encontrava-se muito distante das demais e, portanto, ela foi considerada *outlier* e retirada desta análise. Aqui cabe mencionar que talvez parecesse estranho considerar que tal espécie tenha sido considerada um *outlier* sem que se fizessem neste experimento réplicas em todos os passos da análise, desde o momento da coleta até a injeção dos extratos totais no UHPLC-Orbitrap. Em outras palavras, tal espécie pode de fato ser muito distinta das demais, uma vez que não podemos afirmar com toda a certeza que a sua distinção se deve a algum erro técnico. Ainda assim, trata-se de apenas uma espécie e, para análises supervisionadas cujo intuito seja a classificação e predição de amostras, neste caso é desejável retirar tal amostra e observar como se agrupariam as espécies em uma nova PCA (Figura 17), sem a influência do *outlier*.

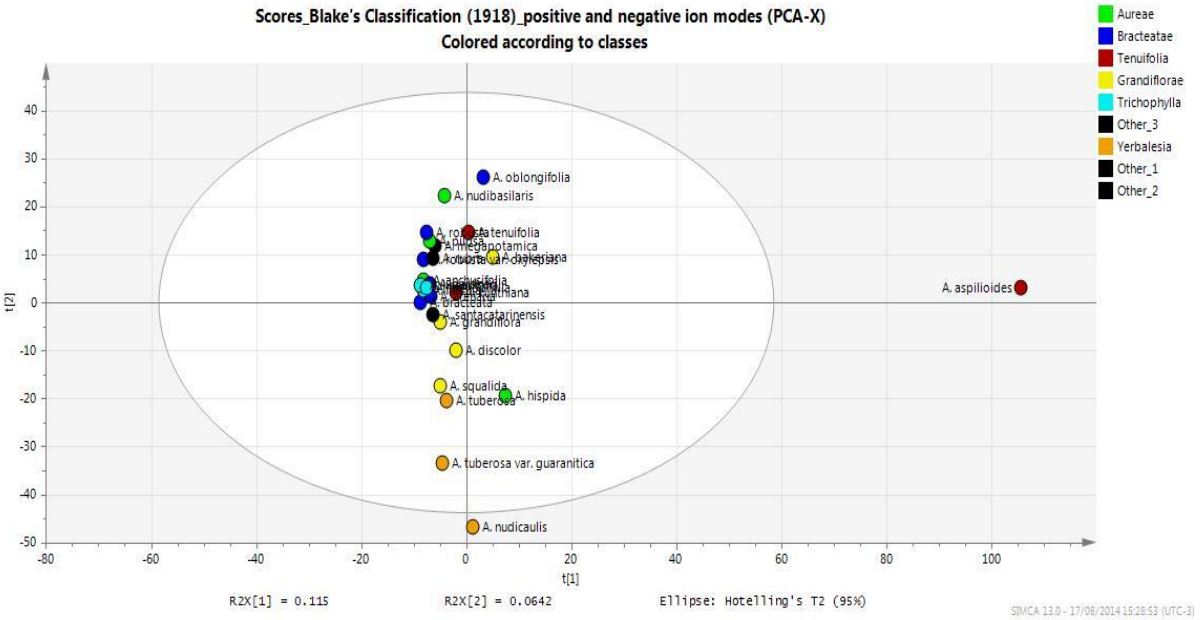


Figura 16 - PCA com base em ambos modos de ionização para o gênero *Aldama* La Llave, de acordo com a classificação infragenérica de Blake (1918) e o *outlier A. aspillioides* (à direita, ao centro).

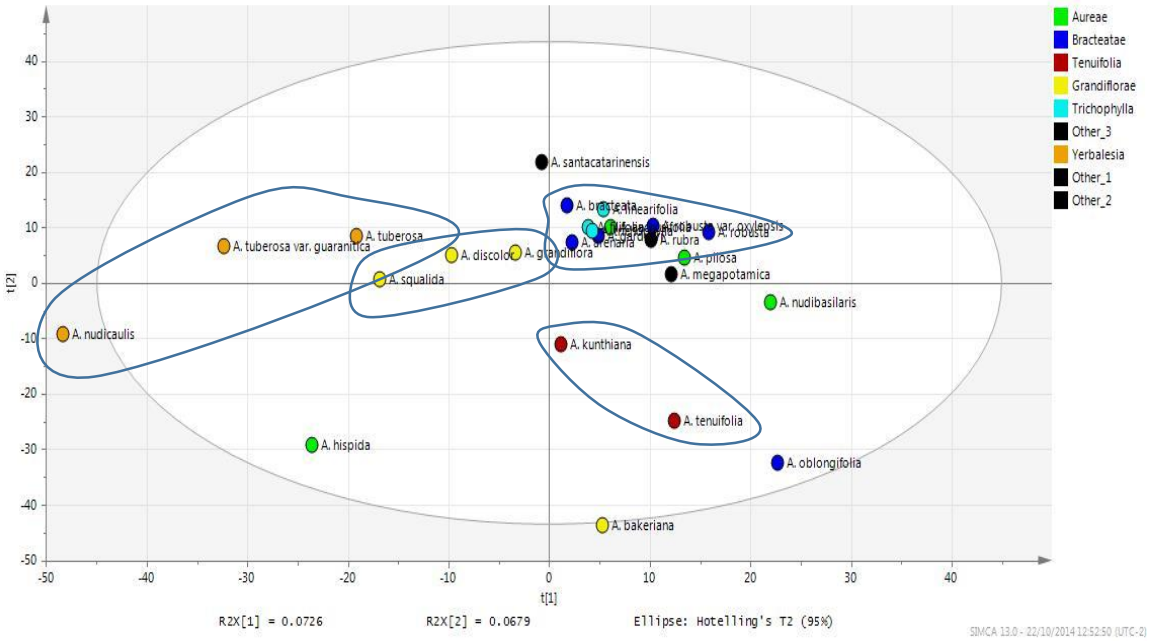


Figura 17 - PCA com base em ambos modos de ionização para o gênero *Aldama* La Llave, de acordo com a classificação infragenérica de Blake (1918) sem a espécie *A. aspillioides*.

Apesar do baixo ajuste observado com apenas duas componentes principais (R^2 = 0,14), poderíamos já notar nesta PCA alguns agrupamentos nas séries Bracteatae, Tenuifoliae, Grandiflorae e para a seção Yearbalesia. Também é evidente que o valor do R^2 não é alto devido ao fato da análise ter sido realizada com as outras três espécies que não tiveram qualquer classificação proposta por Blake. Ainda assim, aqui nós optamos por inseri-las para que pudesse ser visualizado a qual *cluster* em específico elas pertenceriam na PCA, mas isso

não foi possível, visto que elas parecem não pertencer a qualquer grupo distinto. Não obstante, nós realizamos uma nova PCA sem estas três espécies e encontramos um valor de ajuste de R^2 = 0,16, o que também não é muito alto (dados não mostrados). Entretanto, deve-se levar em consideração que estas análises são demasiadas complexas até mesmo para um *software* de quimiometria. Há cerca de 20 mil variáveis (picos detectados) que foram inseridas no *software* para classificar apenas 26 espécies. Isto fatalmente gera um caso complicado para análises multidimensionais e, possivelmente, *overfitting*, que em português se traduz como "sobre ajuste". Para tanto, poderíamos aumentar o número de componentes e checar com quantas seria possível um ajuste de pelo menos 70%, o que é considerado bom em análises metabolômicas. Além disso, poderíamos observar se de fato há *overfitting* no sistema (Figura 18), analisando a correspondência entre ajuste e capacidade preditiva (Q^2).

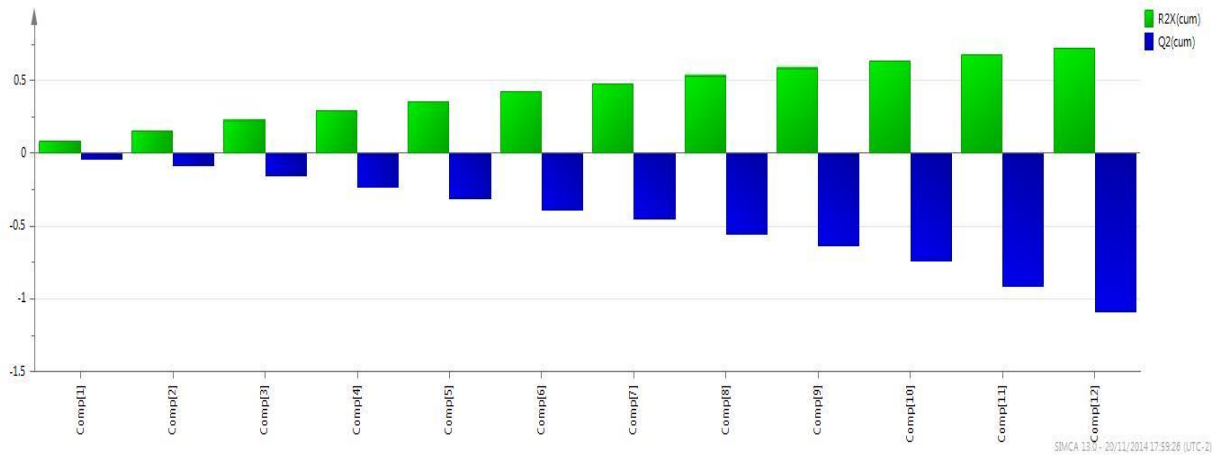


Figura 18 - *Overfitting* na PCA, mostrando o caso em que o ajuste melhora, mas a predição piora.

Assim, pela Figura 18 podemos observar que, à medida em que aumentamos o número de componentes para explicar o modelo o ajuste melhora, porém, a capacidade preditiva piora, o que significa que o modelo está muito ajustado a classificar as amostras, porém, pouco capacitado a realizar predição. O Q^2 é resultante da validação cruzada e nos traz diretamente a informação da quantidade de Y predita por X.

Deste modo, ao se fazer o OPLS-DA, sem *A. aspillioides*, obtivemos uma classificação mais distinta (Figura 19) das seções e séries de Blake. Tal análise foi calculada com cinco componentes preditivas e duas ortogonais, por meio das quais encontramos os melhores valores de ajuste e predição.

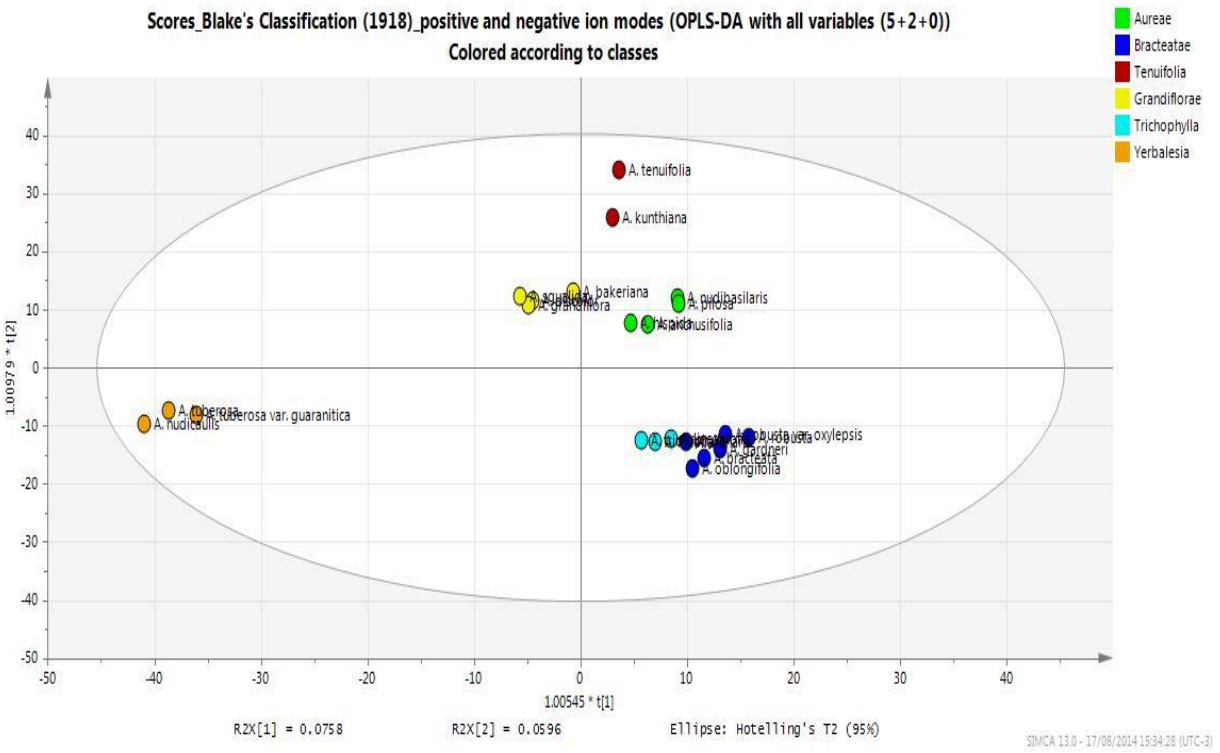


Figura 19 – OPLS-DA com base na classificação de Blake (1918) realizado com ambos os modos de ionização para *Aldama* La Llave, sem *A. aspillioides*.

Entretanto, esta primeira OPLS-DA foi realizada com todas as variáveis químicas (íons detectados) dos extratos totais. Porém por meio do *software* SIMCA foram removidas aquelas variáveis que não são importantes na projeção. Isso é possível porque o *software* fornece uma lista em barras de todas as variáveis da projeção e, ao agrupar tais variáveis em ordem crescente (ou decrescente), pode-se selecionar e remover aquelas que não são importantes para o modelo, tal como mostrado na Figura 20. O processo de remoção de variáveis é sempre realizado aos poucos, cautelosamente, e dentro dos critérios adotados pela literatura (Wold *et al*., 2001; Yuliana *et al*., 2011), a qual preconiza que variáveis importantes são aquelas com um valor acima de 1,0.

Figura 20 - Ilustração da remoção de variáveis não importantes abaixo de 1,0.

Além disso, os modelos gerados por análises metabolômicas trazem uma enorme quantidade de variáveis, da ordem de milhares, sendo que muitas das quais evidentemente não tem qualquer relação com a resposta (matriz Y) que, neste caso, é a classificação de Blake. Assim sendo, quando foram removidas as variáveis não importantes na projeção, obteve-se um agrupamento ainda melhor (Figura 21). Nesta nova análise o número de componentes preditivos e ortogonais foram aumentados para cinco. Importante mencionar aqui que os valores de R^2 das Figuras de OPLS-DA exportadas pelo SIMCA se referem ao valor do ajuste da matriz X explicada. Os valores de R^2 descritos na literatura e que servem de bons parâmetros para averiguar os modelos seriam aqueles que se referem à quantidade da matriz Y explicada por X. Aqui, nos resultados referentes aos estudos quimiotaxonômicos, será usada a mesma notação do SIMCA, que se refere ao ajuste de Y em X como R^2_Y.

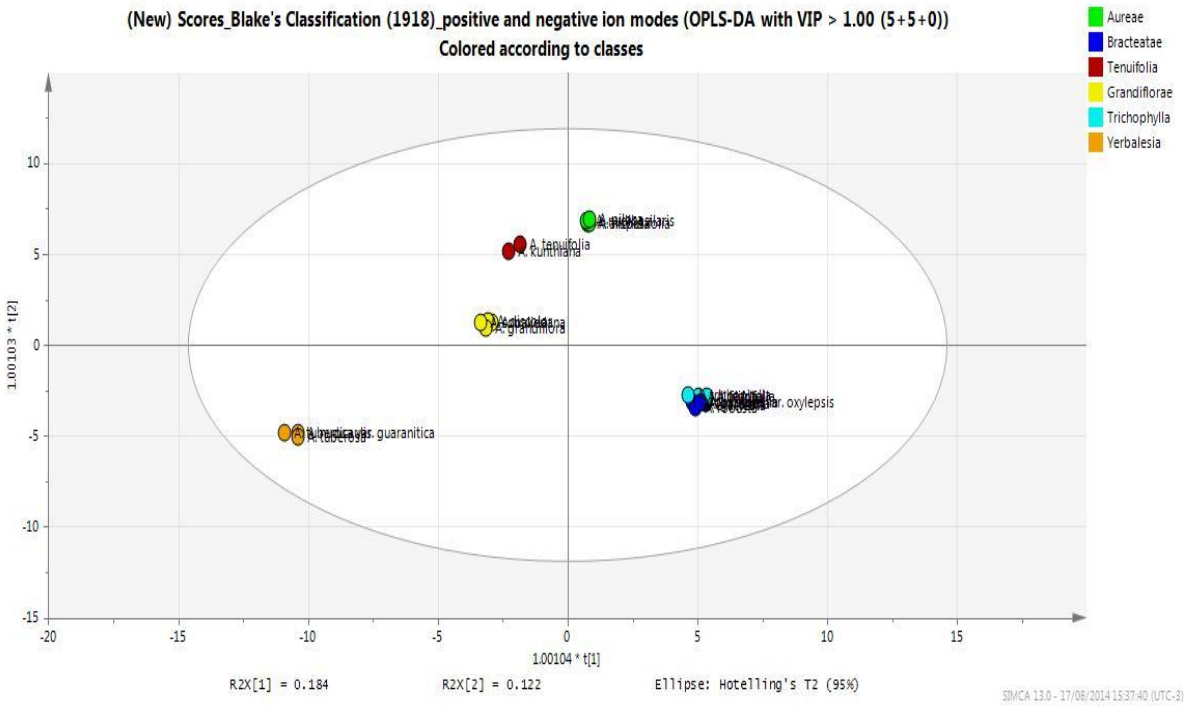


Figura 21 - OPLS-DA otimizada com VIP > 1,0 realizado com base na Classificação de Blake e dados adquiridos por ambos modos de ionização para *Aldama* La Llave.

Os valores de R^2_y e Q^2 desta última OPLS-DA das classificações de Blake estão contidos na Figura 22 exportada pelo *software*. Nela pode-se perceber a contribuição de cada componente, descrita nas linhas horizontais; já nas linhas verticais, os respectivos valores de ajuste e predição. Assim sendo, os valores de R^2_y e Q^2 são 0,99 e 0,87, respectivamente, o que pode ser considerado um ótimo resultado para um modelo de predição. Isso significa que 99% da resposta Y pode ser explicada por X e que 87% pode ser predita.

Workset... Options... Title Blake's Classification (1918)_Aldama La Llave_pos and neg modes

Type: OPLS-DA Observations (N)=22, Variables (K)=349 (X=343, Y=6)

Component	R2X	R2X(cum)	Eigenvalue	R2	R2(cum)	Q2	Limit	Q2(cum)	R2Y	R2Y(cum)	EigenvalueY	Significance
Model		0.76			0.998			0.87		1		
Predictive		0.589			0.998			0.87		1		
P1	0.184	0.184	4.05	0.204	0.204	0.021	0.01	0.021	0.204	0.204	1.22	R1
P2	0.122	0.306	2.69	0.203	0.407	0.0562	0.01	0.0773	0.203	0.407	1.22	R1
P3	0.105	0.411	2.31	0.202	0.608	0.0427	0.01	0.12	0.202	0.609	1.21	R1
P4	0.105	0.516	2.3	0.189	0.797	0.262	0.01	0.382	0.19	0.798	1.14	R1
P5	0.0736	0.589	1.62	0.201	0.998	0.488	0.01	0.87	0.202	1	1.21	R1
Orthogonal in X(OPLS)		0.171			0							
O1	0.0521	0.0521	1.15	0	0							NS
O2	0.0363	0.0884	0.799	0	0							NS
O3	0.0297	0.118	0.653	0	0							NS
O4	0.0255	0.144	0.561	0	0							NS
O5	0.027	0.171	0.595	0	0							NS

Figura 22 - Valores de ajuste e predição alcançados pelo modelo OPLS-DA otimizado (VIP > 1,0).

Tendo em mãos um modelo de excelente capacidade preditiva, foi possível calcular a predição das três espécies que não foram classificadas por Blake (1918). O resultado contido na Figura 23 indica que a espécie *A. megapotâmica* teria 44% de chance em ser classificada como da série Aureae, com base em seus dados químicos. Já as espécies *A. rubra* e *A. santacatarinensis* teriam 37% e 30% de chances de serem classificadas como Bracteatae e Aureae, respectivamente. Os retângulos em vermelho demonstram outras possíveis classificações que se poderia realizar, visto que tais valores são altos e não deveriam ser desprezados. Em outras palavras, evidentemente, uma classificação ideal seria próxima a 100% para uma determinada classe, mas infelizmente isso não foi possível aqui.

Obs ID (Primary)	Obs ID (ClassID)	Pred ((Aureae))	Pred ((Bracteatae))	Pred ((Tenuifolia))	Pred ((Grandiflorae))	Pred ((Trichophylla))	Pred ((Yerbalesia))	TOTAL
A. megapotamica	Other_3	0.44	0.13	0.20	-0.01	0.15	0.10	1.00
A. rubra	Other_1	0.12	0.37	0.04	0.18	0.24	0.05	1.00
A. santacatarinensis	Other_2	0.30	0.22	0.00	0.20	0.05	0.23	1.00

Figura 23 - Valores de predição das três espécies não classificadas por Blake.

Em um modelo de predição, nós certamente também poderíamos correlacionar o gráfico de *scores* do OPLS-DA com seu respectivo gráfico dos *loadings* e, assim sendo, encontrarmos os "porquês" dos agrupamentos e classificações em termos químicos. Entretanto, embora nós

tenhamos gerado uma lista das variáveis discriminantes de cada classe através do SIMCA, nós ainda não pudemos desreplicar tais constituintes químicos visto que, em se tratando de uma abordagem metabolômica, muitos dos metabólitos apontados foram encontrados em baixa intensidade nos TIC e espectros no UV, sendo, portanto, difíceis de serem analisados do ponto de vista do perfil de UV ou dos respectivos padrões de fragmentação. Assim, neste momento caberia uma reflexão sobre a abordagem empregada e, se fosse o caso, alterá-la para uma metabolômica direcionada. Esta ideia será resgatada mais à frente, quando se demonstram os resultados de OPLS-DA para *Aldama* La Llave *versus* os outros gêneros e espécies já mencionados nesta Dissertação.

Por enquanto, apresentaremos os resultados de OPLS-DA com as espécies de *Aldama* classificadas de acordo com as características climáticas do Brasil, uma vez que tais espécies foram coletadas em diversos Estados. Sendo assim, aqui se procura explicar o agrupamento do gênero *Aldama* frente a fatores externos a elas, procurando entender se haveria correlação entre química e as características climáticas, uma hipótese que já foi introduzida e discutida nesta Dissertação. Cabe mencionarmos que tal classificação climática está descrita na Figura 24 que foi publicada para a comunidade científica (Alvares, 2014).

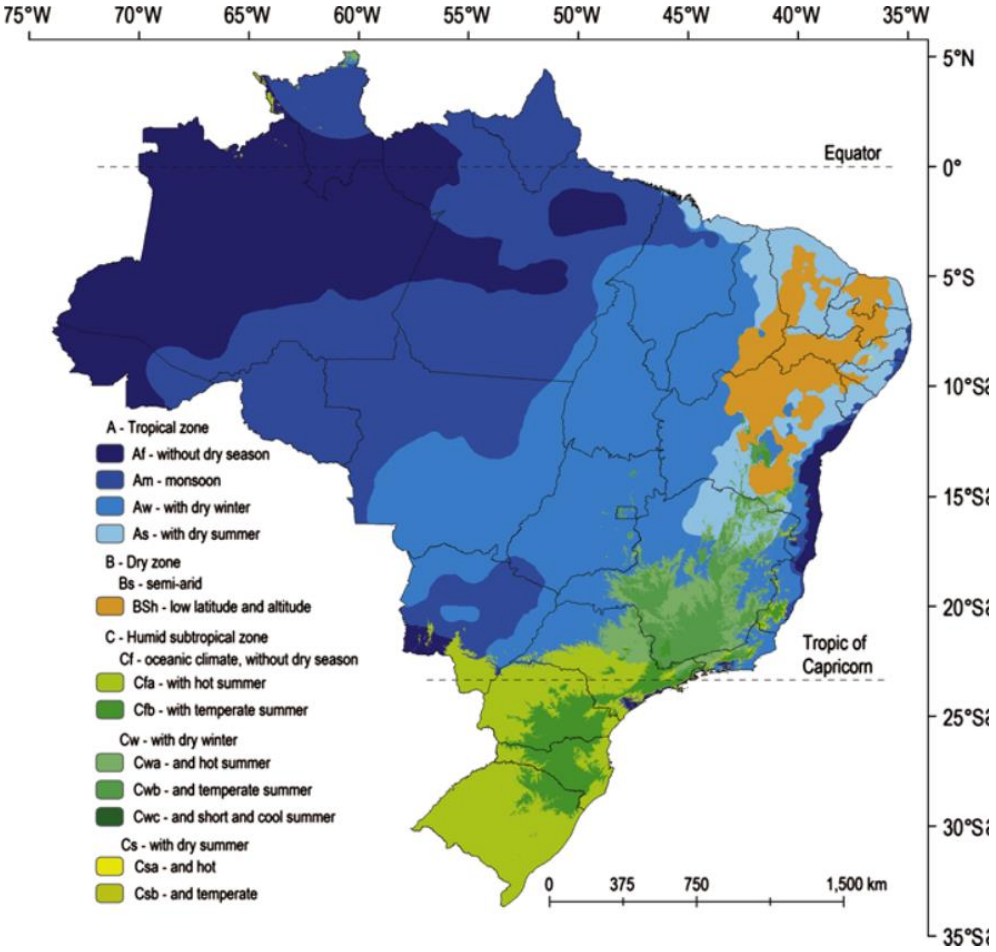


Figura 24 - Classificação climática no Brasil (Extraído de Alvares, 2014).

Deste modo, ao ser utilizado o mapa de Alvares (2014) seria possível localizar por meio das coordenadas geográficas de cada coleta descrita no APÊNDICE 5 (Locais de coleta de *Aldama* La Llave) o local climático em que cada espécie foi coletada e realizarmos o OPLS-DA (Figura 25).

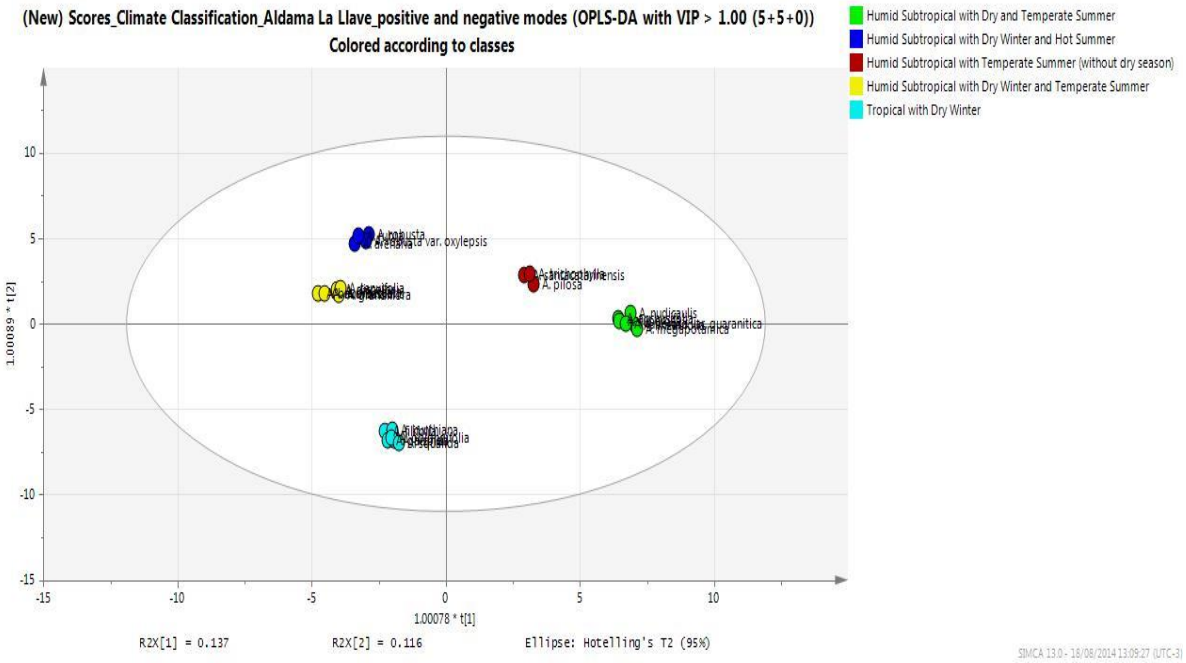


Figura 25 - OPLS-DA com base na classificação climática de Alvares (2014) realizada com dados químicos de *Aldama* La Llave.

É importante informar que tal análise já foi realizada sem a *A. aspillioides*. Além disso, a PCA desta análise não apresentou agrupamentos bem definidos e, consequentemente, foi dificilmente explicada pelos *loadings*. Entretanto, nesta OPLS-DA já com as variáveis não importantes para a projeção retiradas do modelo, foi obtido um R^2_y e um Q^2 de 0,99 e 0,86, respectivamente. Novamente, embora tenha sido possível registrar quais são as variáveis discriminantes, não foi realizado um estudo de desreplicação com elas pelos mesmos motivos acima apresentados. Ainda assim, aqui foi realizada uma outra PCA (Figura 26) a partir desta OPLS-DA, caminho inverso ao que é apresentado comumente pela literatura. Esta PCA não difere de outras, exceto pelo fato dela já não conter as variáveis que não são importantes para o modelo. Em outras palavras, não é uma análise discriminante, tão pouco supervisionada, como o OPLS-DA, mas uma PCA comum.

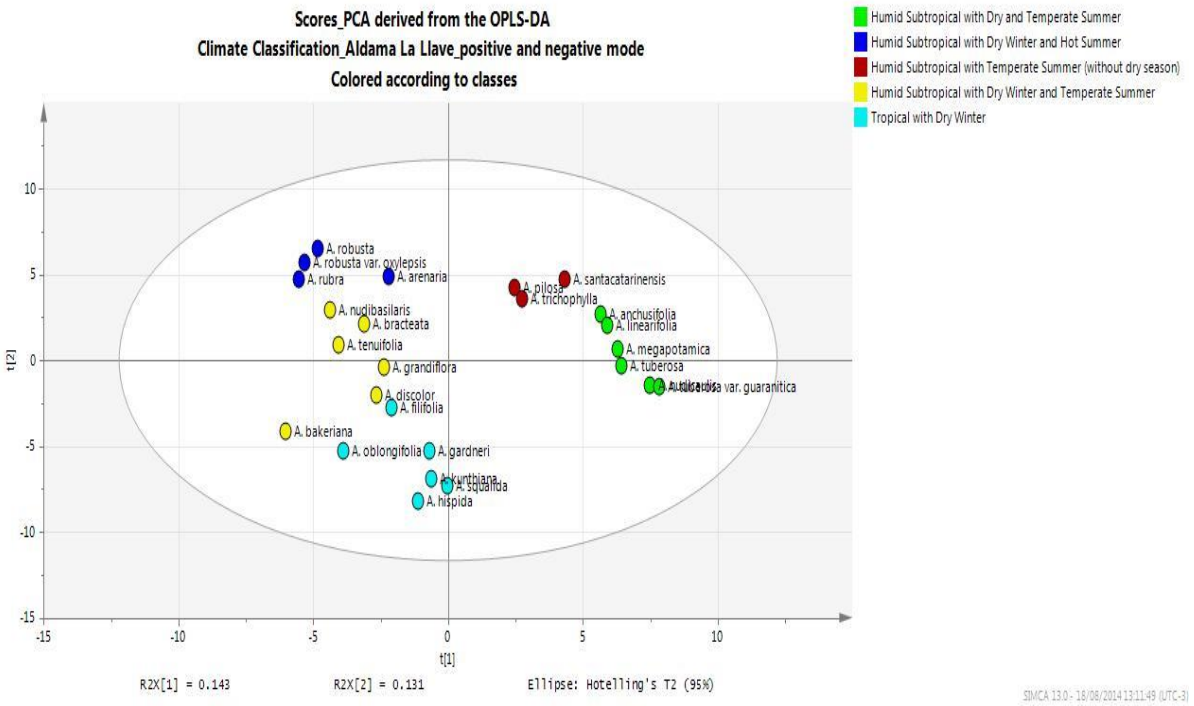


Figura 25 - PCA realizada a partir do OPLS-DA, com VIP > 1,0, para dados químicos de *Aldama* La Llave classificados na classificação climática de Alvares (2014).

Nesta nova PCA, o valor de R^2 é de 0,27, indicando que as duas primeiras componentes explicam 27% do agrupamento observado, de acordo com os dados químicos. Na PCA com todas as variáveis, a qual não está mostrada nesta Dissertação, o valor do ajuste foi de apenas 0,13. Portanto, apenas em se tratando de metabólitos caracterizados como variáveis importantes na projeção, pode-se observar que haveria certa correlação entre o gênero *Aldama* e os diversos locais climáticos. Ainda assim, não se pode assumir aqui que esta seja uma verdade indubitável, uma vez que apenas se presume que tais variáveis não importantes são fragmentos de íons e adutos, os quais não estão contidos nas plantas, pois são oriundos da fonte de ionização empregada, ou ainda poderiam se tratar de metabólitos primários, os quais não são importantes para classificações taxonômicas. Sendo assim, este resultado é mais uma demonstração cautelosa do que uma afirmação categórica daquilo que poderia ser de fato um comportamento real da natureza.

Porém, com as análises de OPLS-DA criadas (Figura 28) com base nas espécies do gênero *Aldama versus* todas as outras espécies e gêneros. O intuito foi separar as *Aldama* em um *cluster* distinto e poder informar quais seriam os *loadings* responsáveis pelo agrupamento, realizando uma posterior desreplicação. Para isto foram utilizadas duas estratégias distintas para a criação de diferentes modelos de predição, visando uma maior chance de encontrar

aquelas que seriam as prováveis substâncias discriminantes. Aqui, uma das estratégias utilizadas não fora apenas a metabolômica não-direcionada, comumente apresentada nesta Dissertação, mas uma outra abordagem semelhante, a qual se utiliza apenas de íons que foram detectados em alta intensidade no EM (acima de 1.499.999 área sob o pico). Com isso esperava-se um maior sucesso no momento da busca de tais íons, encontrando-os e registrando o perfil no UV e padrão de fragmentação do mesmo. Todavia, esta não poderia ser considerada uma abordagem metabolômica-não direcionada de fato, uma vez que já ocorreu um direcionamento para apenas os picos de alta intensidade.

Entretanto, antes de se fazer a OPLS-DA (não mostrada) para este modelo, cabe aqui mencionar que a PCA (Figura 26) realizada com dados de alta intensidade mostra que o gênero *Aldama* ainda não foi misturado aos outros gêneros, revelando que a quimioconsistência do gênero ainda não se perdeu, o que tornaria as variáveis discriminantes geradas pela OPLS-DA mais confiáveis do ponto de vista estatístico. Além disso, a HCA (Figura 27) também revela que o gênero *Aldama* La Llave permanece no mesmo *cluster* com esta nova abordagem.

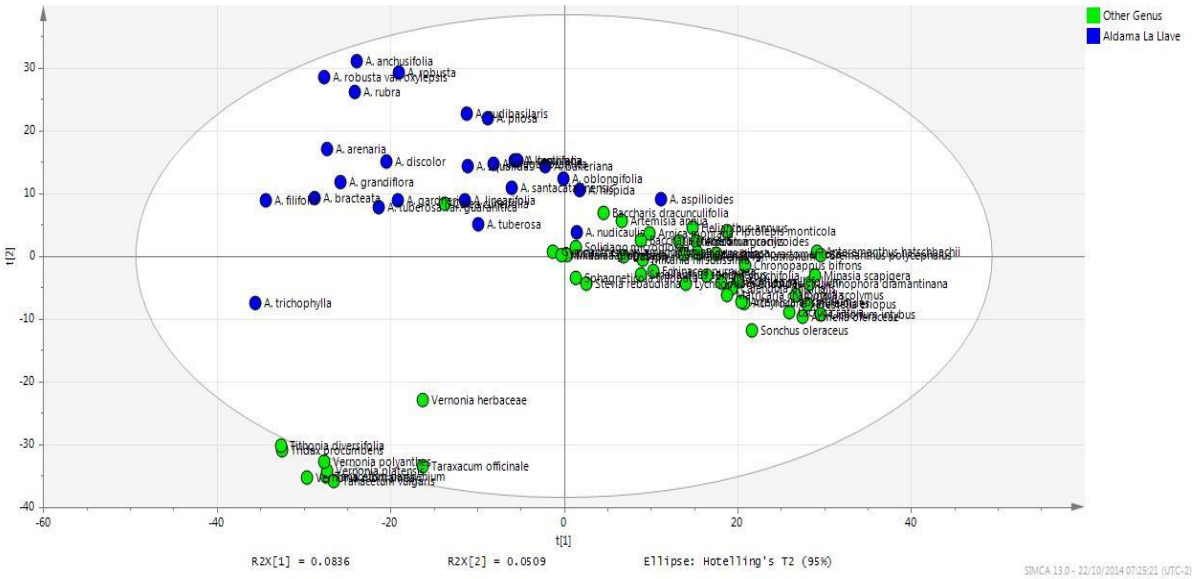


Figura 26 - PCA realizada com picos de alta intensidade detectados por ambos modos de ionização.

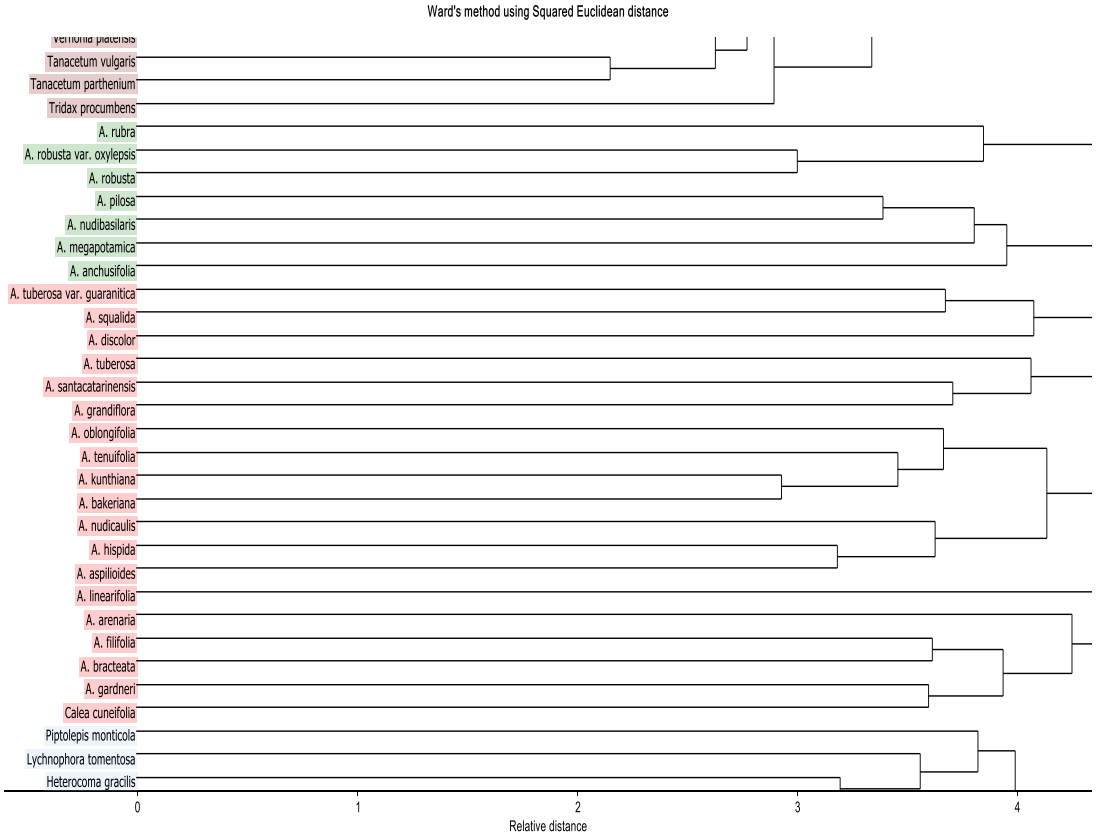


Figura 27 – HCA realizada com picos de alta intensidade detectados por ambos modos de ionização.

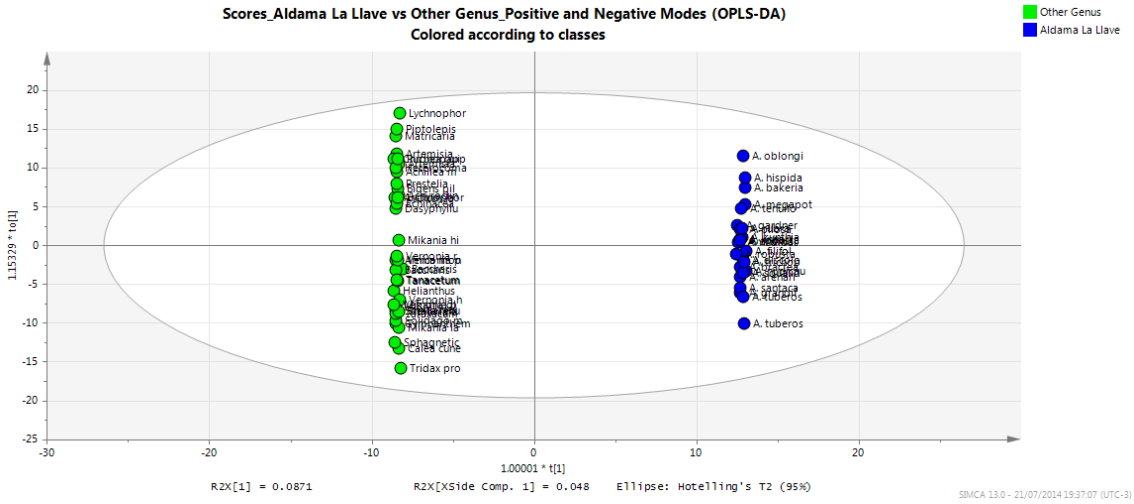


Figura 28 - OPLS-DA realizada com dados provenientes do modo positivo e negativo para o gênero *Aldama* La Llave e outros gêneros.

Assim sendo, com as duas estratégias descritas aqui, obteve-se 66 picos discriminantes apontados pela quimiometria. A *m/z* do íon apresentado foi correlacionada com bancos de AldamaDB e AsteraceaeDB do DNP (*Dictionary of Natural Products*) no próprio *software MZmine*, os quais foram criados com base nas massas das moléculas protonadas,

desprotonadas, possíveis fragmentos e adutos de substâncias de *Aldama* e de outros gêneros de Asteraceae, respectivamente, todas elas já descritas na literatura. Desta forma, chegou-se a 37 prováveis substâncias químicas, as quais deveriam ser desreplicadas com base nos espectros no UV e padrões de fragmentação. Além disso, foi levado em consideração a resolução ($\Delta m/z$) apresentada pelo analisador Orbitrap Exactive Plus, a qual é aproximadamente da ordem de 0,002 de acordo com a literatura (Holcapek *et al*., 2012), para que tivéssemos uma maior certeza no momento de importar as *m*/*z* dos bancos de massas e correlacionar com as massas dos íons detectados. Contudo, infelizmente observou-se que a resolução foi de pouca utilidade no momento da desreplicação. Além disso, mesmo utilizando-se da abordagem com picos de alta intensidade, para auxiliar e se obter maior quantidade de substâncias desreplicadas, muitos espectros de MS-MS apresentaram-se de difícil análise. Por outro lado, não foi observado o espectro de MS-MS de nenhum pico discriminante proveniente da abordagem metabolômica-não direcionada. Assim sendo, neste ponto, tudo o que foi possível foi fazer uma desreplicação parcial das substâncias apontadas pela quimiometria, sugerindo-se algumas prováveis substâncias que estariam mais presentes, ou exclusivas, do gênero *Aldama* La Llave comparado contra todos os outros gêneros. Dados das 37 prováveis substâncias químicas encontradas para o gênero *Aldama* podem ser visualizados na Tabela 1. Algumas substâncias tiveram seus nomes abreviados para STL (lactonas sesquiterpênicas) e *diterpene* (diterpenos), com seus respectivos códigos, uma vez que o nome da substância é muito grande para ser inserido na tabela. Ainda assim, as estruturas químicas destas substâncias podem ser visualizadas na Figura 29.

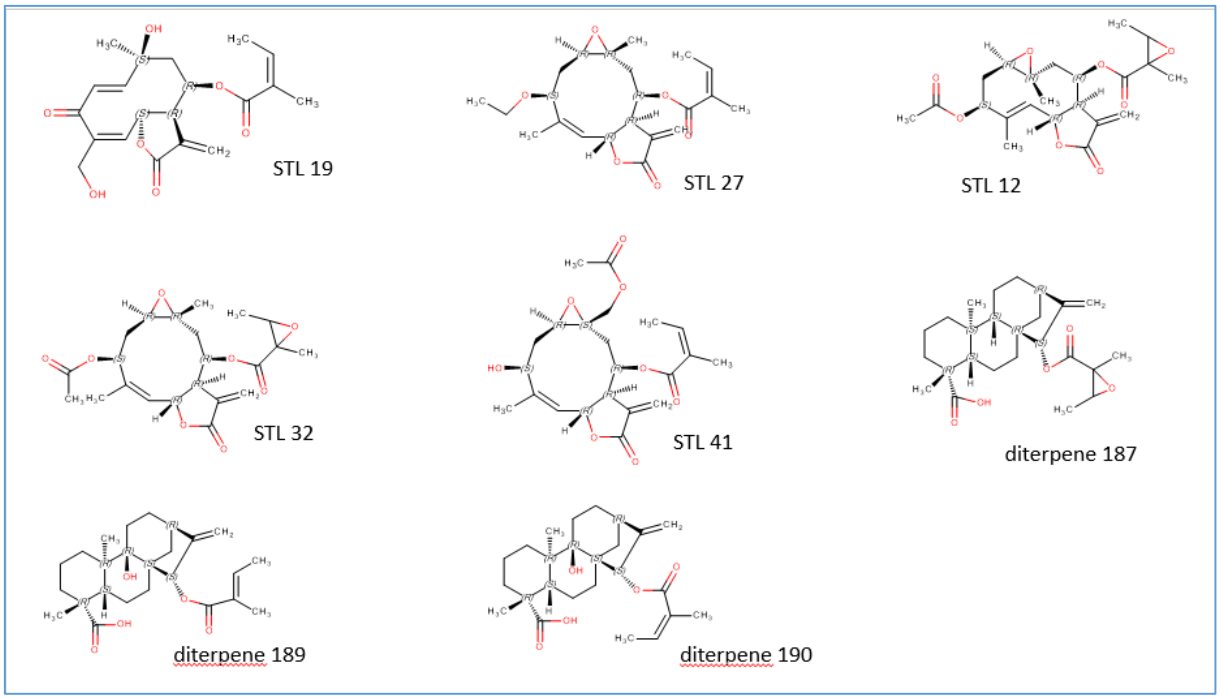


Figura 29 – Estruturas químicas de algumas das prováveis substâncias químicas do gênero *Aldama* La Llave desreplicadas.

Tabela 1 - Lista de prováveis substâncias químicas para *Aldama* La Llave.

ID	Prováveis substâncias do banco (númeração de acordo com o banco AldamaDB) - nomes em inglês
1	coumarinic acid
2	(2*E*)-3-(2,4-dihydroxyphenyl)-1-(4-hydroxyphenyl)prop-2-en-1-one (259)
3	pinocembrin
4	2',4,4'-trihydroxychalcone
5	luteolin (226)
6	kaempferol
7	6-hydroxyluteolin (227)
8	quercetine
9	3',4',7-trimethoxyflavone
10	pilloin (233)
12	ladanein (239)
13	cirsimaritin (241)
14	pectolinarigenin (243)
15	atripliciolide angelate; 15-deoxybudlein A (48)
16	atripliciolide tiglate (49)
17	3-hydroxy-8β-angeloyloxy-1,10-dehydroarbiglovin (99)
18	hymenoxin (246)
20	4',5-dihydroxy-3 3' 6 7-tetramethoxyflavone
21	STL (19)
22	budlein A 2-methylbutyrate; 17,18-dihydrobudlein A (60)
23	budlein A isovalerate (61)
24	4,5-dihydrobudlein A (62)
25	1,2-dehydroniveusin C-2‘,3‘-epoxide (66)
26	1-oxo-1-desoxyniveusin A (74)

Continua na próxima página

27	1,2-anhydroniveusin A (84)
28	3,10α-dihydroxy-8β-angeloyloxyarbiglovin (102)
29	erioflorin acetate; acetylerioflorin (28)
30	STL 059 (27)
31	diterpene (187)
32	diterpene (189)
33	diterpene (190)
34	8β-epoxyangeloyloxy-14-acetoxytithifolin (16)
35	STL 027 (12)
36	STL (32)
37	STL (41)
38	(2S,3S,4S,5R,6S)-6-{[2-(3,4-dihydroxyphenyl)-5,7-dihydroxy-4-oxochromen-3-yl]oxy}-3,4,5-trihydroxyoxane-2-carboxylic acid (271)
39	(2S,3S,4S,5R)-6-{[2-(3,4-dihydroxyphenyl)-5-hydroxy-3-methoxy-4-oxochromen-7-yl]oxy}-3,4,5-trihydroxyoxane-2-carboxylic acid (269)

Além destas substâncias provenientes do banco, o restante dos outros 27 picos apontados pela quimiometria poderão ter sua fórmula molecular calculada com base na *m*/*z* e uma busca por possíveis fórmulas estruturais poderá ser realizada. Além disso, futuramente será realizado um modelo de predição de tempo de retenção construído com base em descritores químicos calculados para 40 estruturas químicas desenhadas em 2D e 3D, e as subst6ancias injetadas nas mesmas condições analíticas do experimento apresentado nesta Dissertação, visando-se predizer o tempo de retenção das estruturas químicas apontadas pela quimiometria. Tal abordagem que utiliza a predição de tempos de retenção tem sido explorada pela academia (Henchoz *et al*, 2009; Eugster *et al*., 2014) e, em nossa equipe científica, foi tema de projeto de doutorado desenvolvido no exterior (Universidade de Münster – Alemanha) pelo aluno Tiago Branquinho de Oliveira (processo FAPESP nº 2013/14239-5).

Por fim, tal como mencionado em "Material e Métodos", fora realizada a extração por lavagem em EtOH (ELF) e tricomas glandulares (TG) das folhas. Estes extratos também foram analisados via UHPLC-PDA-Orbitrap e comparados por meio das análises multidimensionais. Inicialmente, a ideia aqui era explorar quimicamente cada distinto grupo formado pela OPLS-DA, entretanto, tendo em vista as dificuldades apontadas em relação à desreplicação, foi realizada apenas a classificação evidenciando que a mesma poderia ser utilizada como um método analítico e, ainda, preditivo, por meio do qual novos extratos poderiam ser classificados ou preditos. Deste modo, foi realizada uma PCA com os ET, ELF, ELI, TG e TGI, onde se observou que os TG são mais distintos entre si, revelando inclusive espécies muito distintas do ponto de vista químico, tais como *A. bakeriana* e *A. grandiflora*. Assim sendo, estas espécies foram retiradas do modelo, visando otimizá-lo (dados não mostrados). A partir disto, a primeira OPLS-DA realizada também foi otimizada, excluindo-se as variáveis importantes para a projeção com valor inferior a 1,0 (VIP > 1,0). Estes passos resultaram no modelo para predição de tipos de extração (*Extraction type*) representado pela Figura 30.

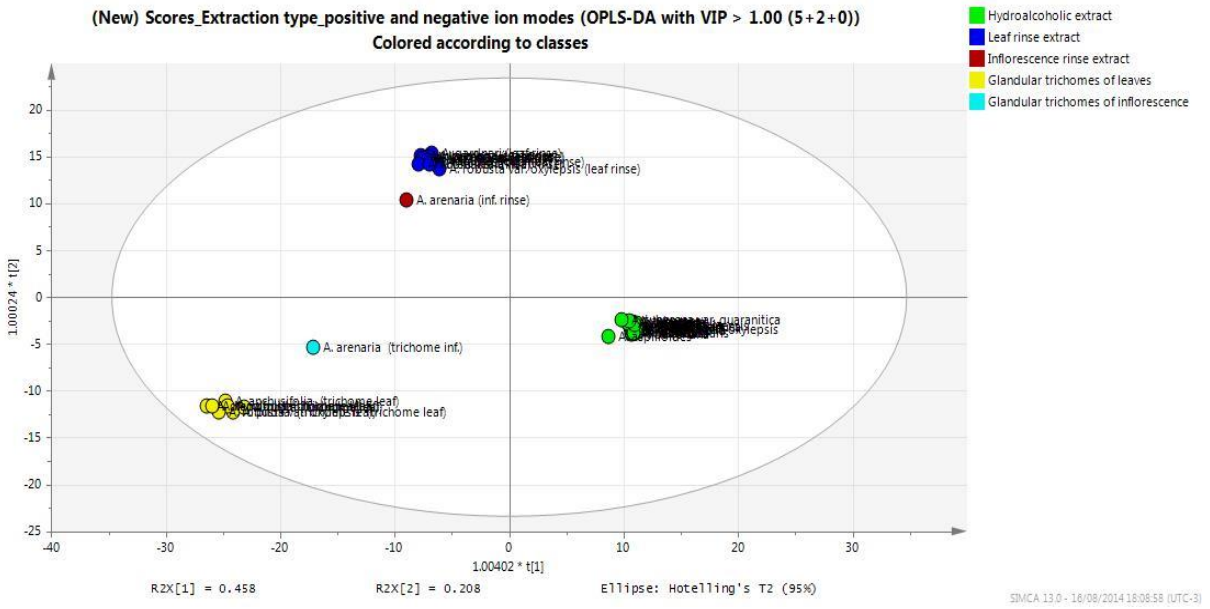


Figura 30 - OPLS-DA para modelo de predição de tipos de extração

Neste modelo pode-se perceber que, certamente, há constituintes químicos que estejam aproximando as extrações realizadas com as inflorescências de *A. arenaria*, os quais são possivelmente LST específicas de inflorescências, uma vez que tanto a extração de TG, quanto a lavagem – de folhas ou inflorescências – com EtOH são técnicas utilizadas para se obter extratos ricos em LST. Estes grupos bem distintos observados no OPLS-DA já haviam se formado pela PCA. Este é um indicativo de que a extração seletiva é uma boa alternativa para se trabalhar com metabolômica direcionada, por meio da qual se poderia desreplicar classe específicas de substâncias químicas. Por fim, o ajuste e a capacidade preditiva apresentados por este modelo foi de $R^2_y = 0,99$ e $Q^2 = 0,94$, respectivamente.

Resultados e Discussão dos estudos de inibição dupla das enzimas COX-1 e 5-LOX

3 – Resultados dos ensaios de inibição enzimática COX-1 e 5-LOX & Comparação dos diferentes modelos criados.

A Tabela 2 mostra os resultados referentes aos ensaios de inibição das enzimas COX-1 e 5-LOX realizados com extratos totais de 23 espécies de *Aldama*. Pode-se verificar que apenas duas espécies demonstraram satisfatoriamente capacidade de inibição de ambas enzimas, tal como idealizado.

Tabela 2 - Resultado de inibição das enzimas COX-1 e 5-LOX para *Aldama* La Llave.

Resultado de inibição das enzimas COX-1 e/ou 5-LOX			
ID	**Espécie**	**COX**	**LOX**
1	*A.anchusifolia* (DC.) Baker	Sem atividade	41,08 µg/mL
2	*A.arenaria* Baker	Sem atividade	16,50 µg/mL
3	*A.bakeriana* S.F.Blake	0,1 µg/mL	Sem atividade
4	*A.bracteata* Gardner	Sem atividade	3,60 µg/mL
5	*A.discolor* Baker	Sem atividade	8,40 µg/mL
6	*A.filifolia* Sch.Bip. ex Baker	Sem atividade	Não
7	*A.gardneri* Baker	Sem atividade	Sem atividade
8	*A.gardneri* Baker var. *gardneri*	Sem atividade	80,42 µg/mL
9	*A.grandiflora* (Gardner) Gardner	Sem atividade	71,06 µg/mL
10	*A.kunthiana* Gardner	Sem atividade	15,12 µg/mL
11	*A.linearifolia* Chodat & Hassl.	Sem atividade	Sem atividade
12	*A.megapotamica* Malme	Sem atividade	68,17 µg/mL
13	*A.nudibasilaris* S.F.Blake	Sem atividade	67,60 µg/mL
14	*A.nudicaulis* Baker	Sem atividade	Sem atividade
15	*A.pilosa* Baker	> 1000 µg/mL	67,22 µg/mL
16	***A.robusta* Gardner**	**0,1 µg/mL**	**36 µg/mL**
17	*A.robusta* var. *oxylepsis* S.F.Blake	Sem atividade	108,70 µg/mL
18	*A.rubra* Magenta	Sem atividade	0,78 µg/mL
19	*A.squalida* S.Moore	Sem atividade	Sem atividade
20	*A.subdentata* S.F.Blake	Sem atividade	51,39 µg/mL
21	*A.tenuifolia* Gardner	Sem atividade	>100 µg/mL
22	***A.trichophylla* Dusén**	**2,6 µg/mL**	**12 µg/mL**
23	*A.tuberosa* Griseb.var. *guaranitica*	Sem atividade	Sem atividade

Estas 17 espécies analisadas nos ensaios *in vitro* foram analisadas via HPLC-PDA-TOF e UHPLC-PDA-Orbitrap, de acordo como descrito em “Material e Métodos”, por meio dos quais se obteve os dados químicos que foram posteriormente tratados e processados por meio dos *softwares MZmine*, *MetAlign* e *MSClust*. Assim sendo, com os dados químicos gerados das 32 combinações criadas (16 provenientes de cada modo de ionização) a partir dos

diferentes algoritmos e parâmetros do *MZmine* foram realizados os modelos de predição de inibição de COX-1 e 5-LOX no *software* SIMCA através de análises OPLS-DA. Deste modo, as Figuras 31 e 32 nos mostram os resultados obtidos para os valores de ajuste (R^2 = quantidade de Y explicados pelo modelo) alcançados com os modelos realizados com dados do modo de ionização positivo do TOF e Orbitrap, respectivamente.

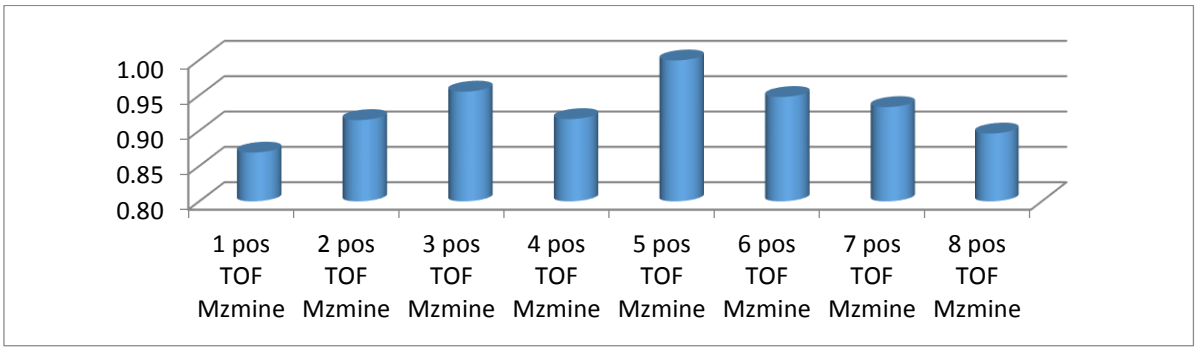


Figura 31 – Valores de ajuste alcançados para dados provenientes do TOF e tratados por diferentes algoritmos do *MZmine.*

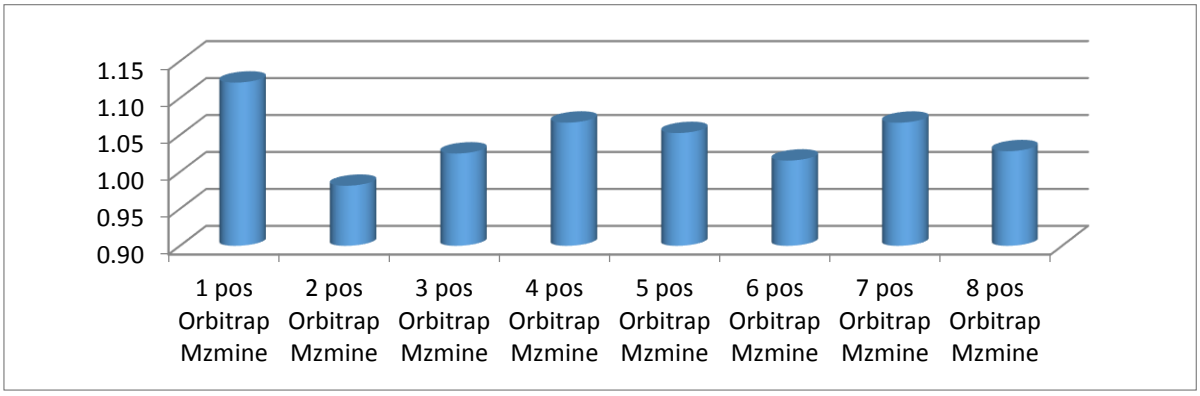


Figura 32 - Valores de ajuste alcançados para dados provenientes do OBT e tratados por diferentes algoritmos do *MZmine.*

A informação que podemos prontamente extrair das Figuras 31 e 32 é que não existe relação direta entre um determinado algoritmo escolhido no *MZmine* e uma melhora no ajuste. Outra evidencia obtida é a de que, em termos de ajuste, todos os modelos são bons. Todavia este resultado não poderia ser considerado surpreendente, uma vez que não é trabalho difícil para o OPLS-DA encontrar em um espaço multidimensional uma ótima separação entre apenas duas distintas classes, tais como “COX-1 e 5-LOX” (inibição dupla) *versus* “5-LOX”

(Tellinghuisen e Bolster, 2011). Portanto, uma boa forma de se avaliar um modelo certamente é por meio de sua capacidade preditiva (Westerhuis, *et al*., 2008). Assim sendo, o Gráfico 1 mostra exatamente os valores da capacidade preditiva de cada modelo criado.

Primeiramente, o que podemos perceber é que, em termos de capacidade preditiva, estes resultados não poderiam ser considerados como bons, mesmo para estudos metabolômicos. Na verdade, valores de Q^2 abaixo de 0.25 são, de fato, muito baixos. Ainda assim, poderíamos relevar tal situação se considerássemos que tais modelos podem ser otimizados se fossem removidas as variáveis não importantes na projeção, o que será discutido mais adiante. Outro fato que podemos observar é que parece que os resultados de Q^2 encontrados foram menos piores para os modelos realizados com dados provenientes do HPLC-TOF, do que com os dados do UHPLC-Orbitrap. Além disso, os dados que foram adquiridos pela ionização no modo negativo também aparentam ser menos piores.

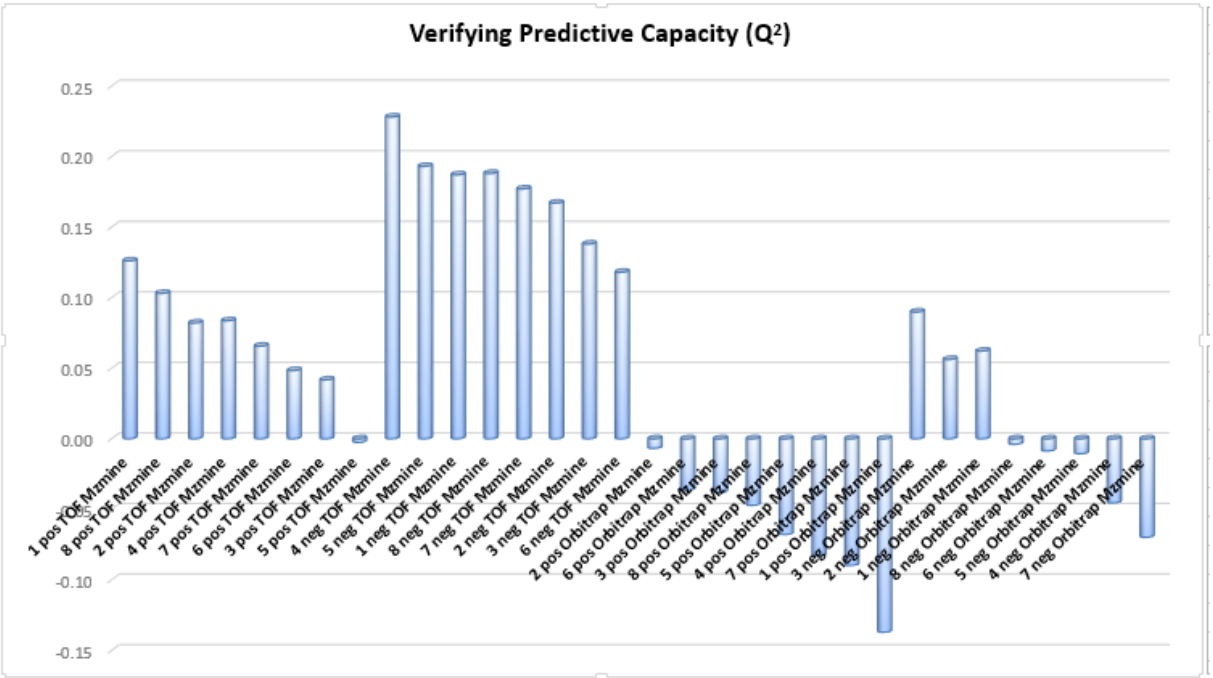


Gráfico 1 - Capacidade preditiva de cada modelo criado com dados tratados via *MZmine.*

A partir destes resultados foi escolhido o melhor modelo criado com dados provenientes do modo de ionização positivo e negativo tanto do HPLC-TOF quanto do UHPLC-Orbitrap, totalizando quatro modelos. Estas quatro melhores combinações do *MZmine* foram processadas *a posteriori* com o *software MSClust*, com a finalidade de verificar se obteríamos

uma melhora dos modelos de predição, uma vez que o *MSClust* agrupa fragmentos ou adutos de íons em seus respectivos prováveis metabólitos, os quais são provenientes das análises ESI-MS e, portanto, não estão presentes no momento dos ensaios *in vitro*. Além disso, estas duas técnicas (tanto o *MZmine* operando sozinho, como em parceria com o *MSClust*) foram comparadas com os dados químicos provenientes dos modos de ionização positivo e negativo provenientes do HPLC-TOF e UHPLC-Orbitrap processados pelo *MetAlign* em conjunto com o *MSClust*.

Destas comparações resultou o Gráfico 2, no qual pode-se notar que não parece haver melhora significativa quando há utilização do *MSClust* com dados provenientes do *MZmine*. De fato, o *MSClust* agrupa um conjunto de picos em um "metabólito putativo", o qual pode ser íons precursores, adutos, fragmentos, dímeros etc., não necessariamente um íon precursor. A principal tarefa deste *software* é reduzir os dados para um conjunto menor de variáveis, que representam todas as informações presentes em um arquivo de entrada.

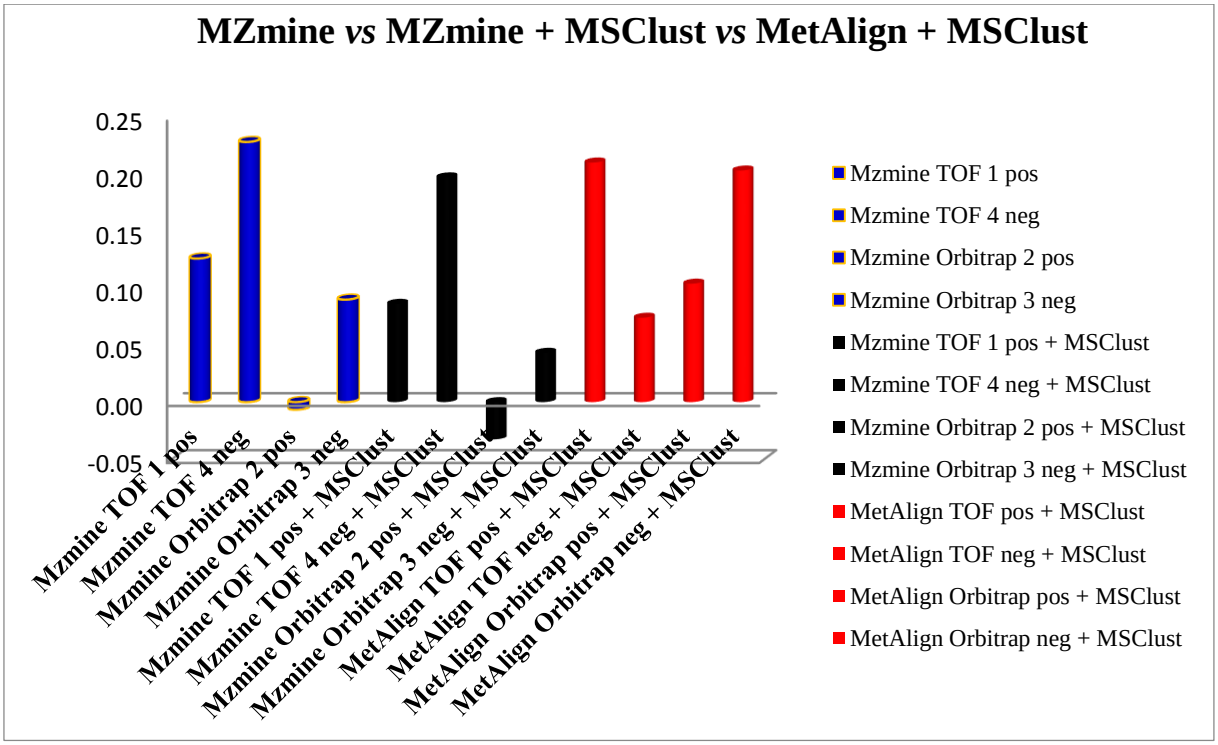


Gráfico 2 - Comparação dos diferentes softwares de tratamento utilizados para dados provenientes do modo positivo e negativo do TOF e OBT.

Por outro lado, com os dados tratados com os *softwares MetAlign* e *MSClust* em conjunto não foi possível encontrar uma melhora significativa nos modelos, tão pouco fora encontrada uma correlação entre a utilização deles com qualquer analisador de massas ou modo de ionização. Assim sendo, destes 12 modelos criados, foram escolhidos os sete mais

promissores (Gráfico 3) para a etapa de otimização (próximo item), com deleção de variáveis não importantes e inserção de componentes ortogonais ao modelo de predição de inibição de COX-1 e 5-LOX via OPLS-DA.

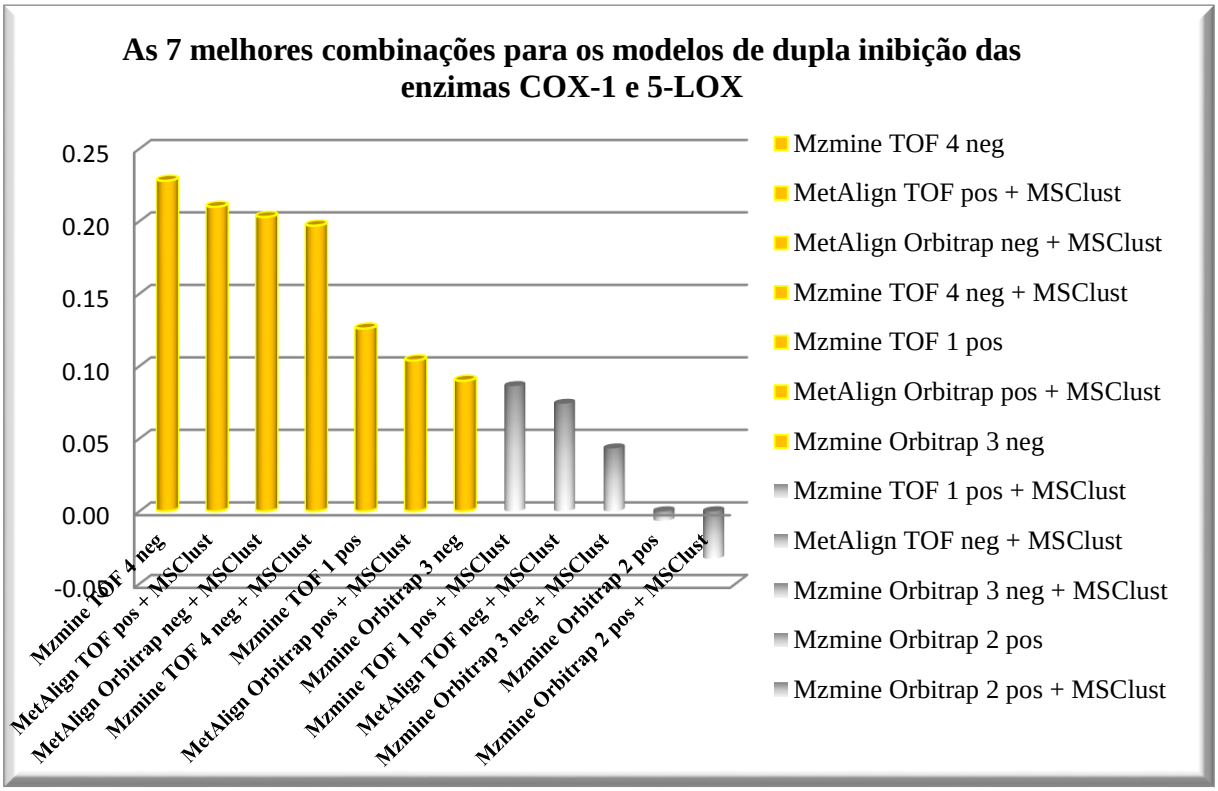


Gráfico 3 – Melhores combinações para todos os modelos criados com base em dados químicos provenientes dos modos positivo e negativo do TOF e OBT.

4 – Otimização dos melhores modelos de inibição enzimática & Desreplicação direcionada.

Após terem sidos selecionados os sete melhores modelos, estes foram otimizados eliminando-se as variáveis não importantes para a projeção do OPLS-DA. Neste caso, supõe-se que tais variáveis não importantes sejam fragmentos de íons, os quais não estão presentes em extratos analisados nos ensaios *in vitro*, ou ainda metabólitos primários que podem ter sido extraídos e estejam sendo detectados pelo espectrômetro de massas. Assim sendo, foi gerado o Gráfico 4 que demonstra, por meio da comparação dos valores de Q^2 com até duas componentes ortogonais e uma componente preditiva, a melhora significativa que ocorre nos modelos de predição via OPLS-DA após a deleção de variáveis não importantes.

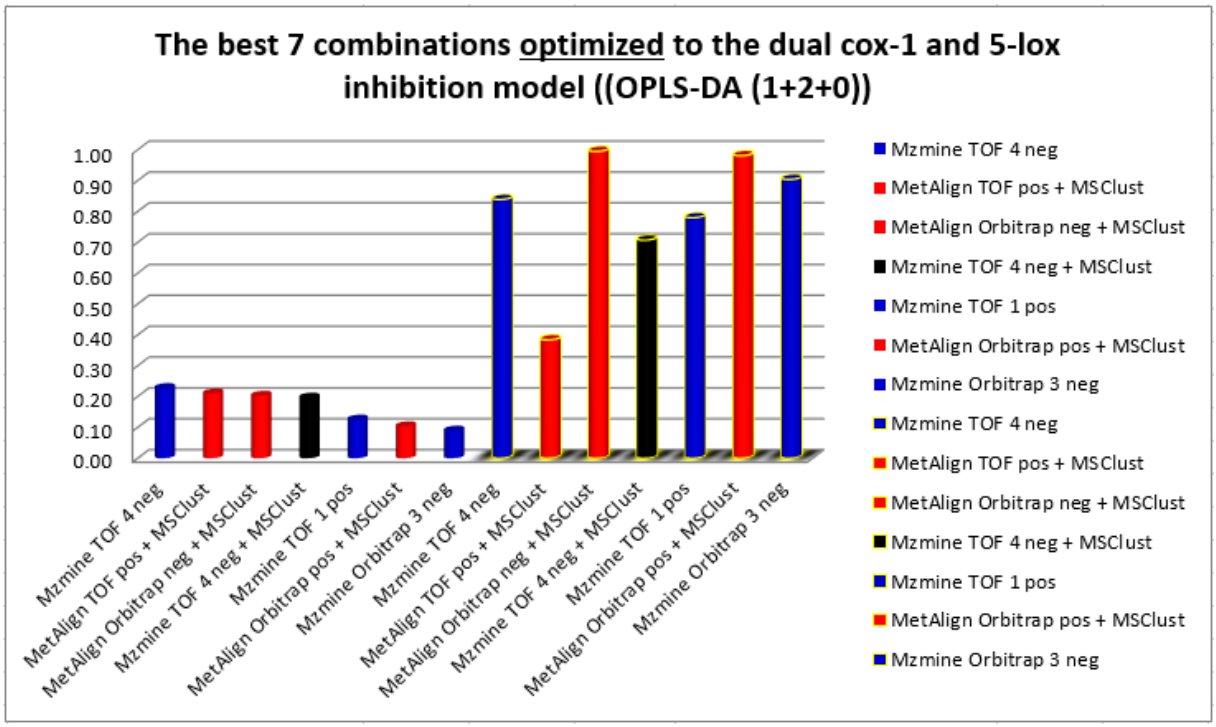


Gráfico 4 - As melhores combinações otimizadas por meio da deleção de variáveis não importantes para a projeção (VIP < 1,0).

Outra forma de melhorar o modelo é aumentar o número de componentes preditivas e/ou ortogonais. Esta prática também é comum e permitida através do *software* SIMCA, desde que o processo não ocasione *overfitting*, o que pode ser percebido quando o valor de R^2 aumenta e, ao mesmo tempo, ocorre uma queda no valor de Q^2. O aumento do valor do ajuste do modelo deve sempre acompanhar o aumento do valor da capacidade preditiva do modelo, sendo que, de forma contrária, indicaria *overfitting* (Figura 33 – retirada do Manual do SIMCA 13.0.3).

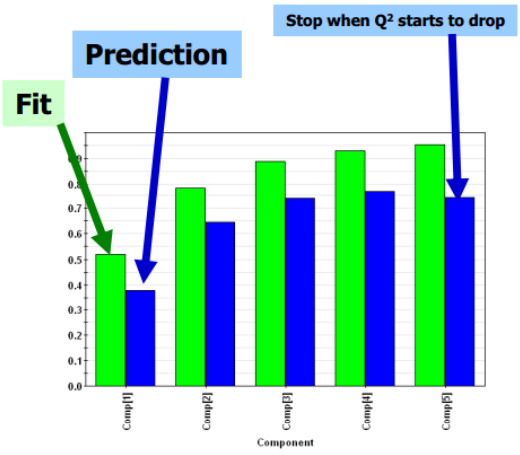


Figura 33 – Indicação de *overfitting* (Extraída do Manual do SIMCA 13.0.3).

Assim sendo, os seis melhores modelos (desta vez, excluindo-se o modelo do *MetAlign* com dados do TOF modo positivo), apresentam uma ligeira melhora quando utilizamos o máximo de componentes possíveis para explicar tal modelo, demonstrado no Quadro 2. Tal quadro mostra os valores de Q^2 com apenas 2 componentes ortogonais ((1+2+0) e os valores com um número máximo permitido de componentes principais.

	(1+2+0)		
Model	**Q^2**	**Q^2**	**máximo nº ortogonais permitido**
Mzmine TOF 4 neg	0.84	0.92	6
MetAlign Orbitrap neg + MSClust	1.00	1.00	2
Mzmine TOF 4 neg + MSClust	0.71	0.79	6
Mzmine TOF 1 pos	0.78	0.87	5
MetAlign Orbitrap pos + MSClust	0.98	0.99	3
Mzmine Orbitrap 3 neg	0.90	0.97	5

Quadro 2 – Modelos otimizados com o máximo número de componentes ortogonais permitido.

A partir destes seis melhores modelos otimizados, cada qual pode ser utilizado para se fazer a análise discriminante, com o intuito de gerar as variáveis (substâncias químicas) discriminantes (mais importantes) na inibição dupla das enzimas COX-1 e 5-LOX. De tal maneira, pode-se observar no gráfico do OPLS-DA, gerado com os dados químicos provenientes do modo negativo do Orbitrap e tratados pela combinação 3 do *MZmine 2.9* (MZmine Orbitrap 3 neg), que os extratos ativos contra tais enzimas pró-inflamatórias (em azul) foram separados dos demais extratos (em verde), os quais são ativos apenas contra 5-LOX (Figura 34).

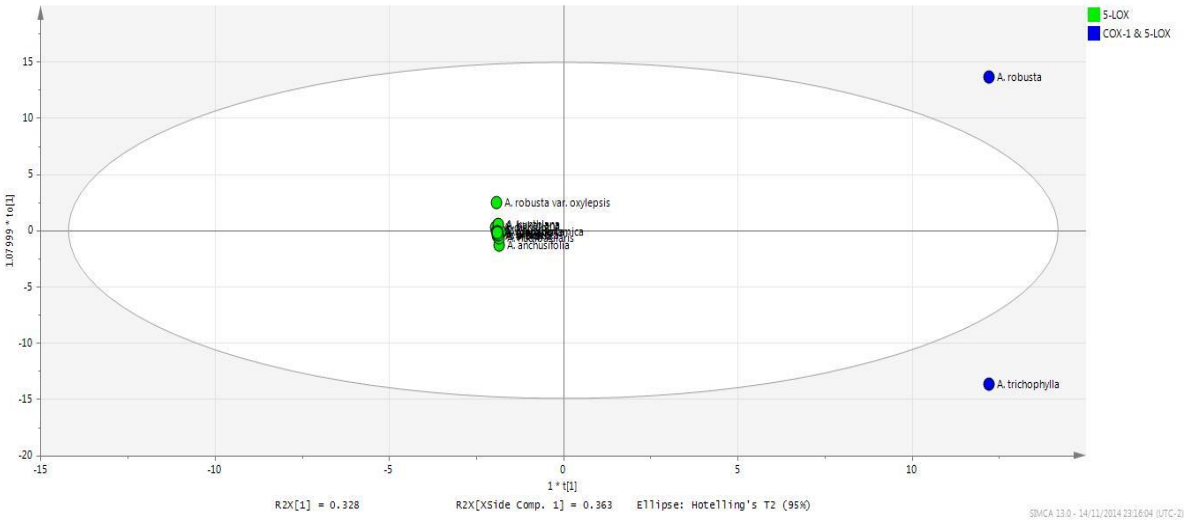


Figura 34 – Modelo de inibição enzimática por OPLS-DA criado com dados provenientes do modo negativo do Orbitrap e processados pelo MZmine (3 neg OBT).

Na Figura 34, pode-se notar que, embora os extratos ativos contra COX-1 e LOX-5 tenham se separado dos demais extratos, eles não são muito semelhantes do ponto de vista químico, visto que eles se encontram muito separados. Ainda assim, é de se esperar que as substâncias químicas compartilhadas pelas duas espécies sejam as inibidoras das enzimas inflamatórias.

É importante mencionar que todos os outros modelos geraram gráficos semelhantes ao da Figura 34, embora cada qual tenha tido uma capacidade preditiva diferente. Além disso, o *software* SIMCA permite analisar os valores de Raiz Média do Erro Quadrático da Validação-Cruzada (*Root Mean Square Error of Cross-Validation* – RMSEcv), os quais devem ter o menor valor possível mensurados na mesma magnitude da classe analisada no modelo OPLS-DA. Melhor explicando, neste estudo a classe analisada é nominal, pois não foi informado o valor numérico da concentração inibitória das enzimas, mas apenas se inibiu (valor = 1) ou não (valor = 0).

O RMSEcv fornece os valores do erro da validação-cruzada, a qual é utilizada para gerar o modelo OPLS-DA. O *software* SIMCA faz a validação-cruzada utilizando-se do *7 fold cross-validation leave-one-out*, o que significa que ele, primeiramente, particiona o espaço amostral em sete, e retira a sétima parte para testar a capacidade preditiva do modelo gerado pelas outras seis partes juntas, repetindo o processo por sete vezes com diferentes amostras.

Em princípio, o RMSEcv cai à medida que se utiliza maior número de componentes (preditivas ou ortogonais) para se realizar um modelo, o que pode ser observado na Figura 35 proveniente da mesma análise dos dados do “MZmine Orbitrap 3 neg”.

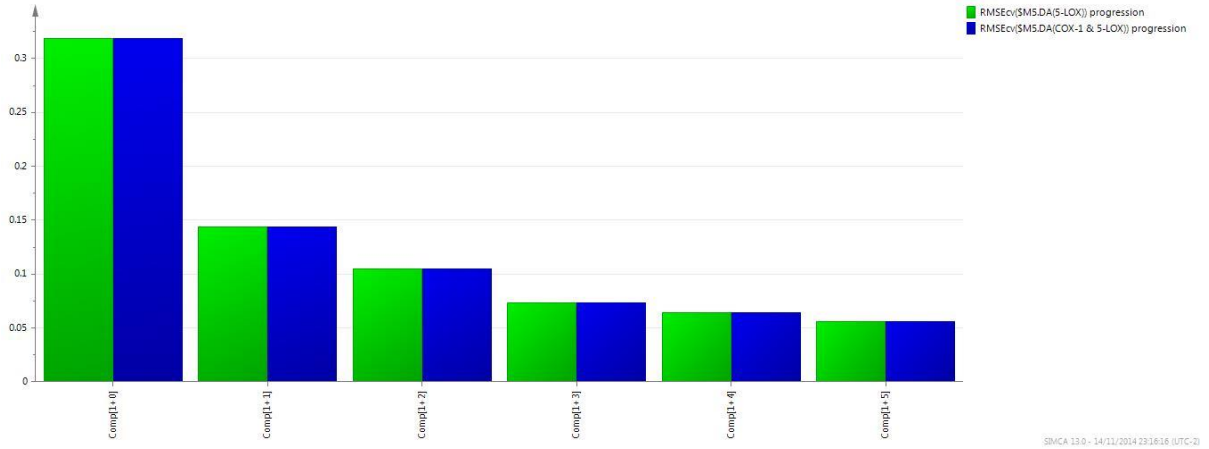


Figura 35 – RMSEcv em decaimento com o aumento do número de componentes do modelo de predição

Como a variável é nominal, assumido valor máximo igual a 1, o modelo pode ser considerado muito bom, visto que o RMSEcv se encontra em praticamente 0 quando utilizamos cinco componentes, para este modelo. Os outros valores de RMSEcv dos outros modelos serão mostrados mais adiante.

Tendo em mãos bons modelos preditivos, podemos gerar as variáveis mais importantes, ou discriminantes, na inibição de COX-1 e 5-LOX. Entretanto, vale mencionar aqui que neste projeto de pesquisa não apenas se preocupou em gerar as prováveis substâncias químicas importantes na inibição de COX-1 e 5-LOX, mas em desreplicá-las, isolá-las e, se possível, em testá-las em ensaios *in vitro*, confirmando a atividade, tal como descrito em Objetivos. Deste modo, apenas a espécie *A. trichophylla* foi escolhida para este intuito. Portanto, a variável discriminante deve ser não apenas inibidora de COX-1 e 5-LOX, mas estar presente em *A. trichophylla*. Felizmente, podemos fazer isso por meio do SIMCA, quando primeiramente selecionamos a espécie de interesse e pesquisamos quais são as variáveis importantes na sua contribuição (*Contribution Plot*). Podemos ver este processo na Figura 36 do OPLS-DA do modelo “MZmine TOF 1 pos”, quando a *A. trichophylla* está selecionada (em vermelho) e geramos o gráfico de *Contribution Plot*.

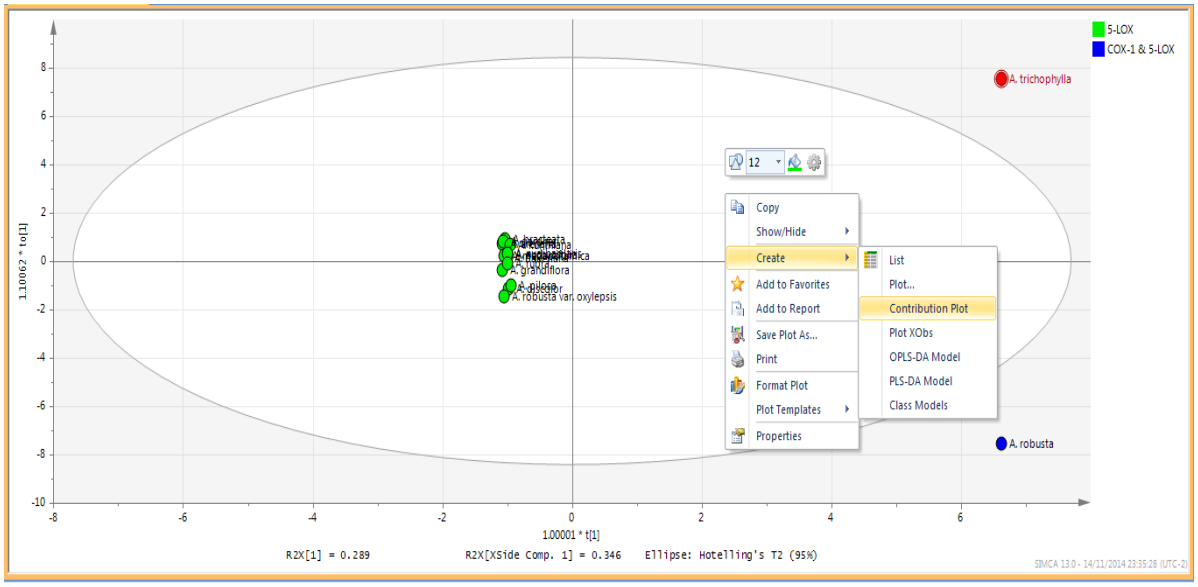


Figura 36 - Destaque para o processo que gera o *Contribution Plot* para *A. trichophylla*.

Ao clicarmos em *Contribution Plot* o gráfico de variáveis importantes para *A. trichophylla* será gerado (Figura 37), evidenciando no eixo das abscissas os valores de massa das moléculas protonados detectadas pelo TOF no modo positivo e processados pela combinação 1 do MZmine.

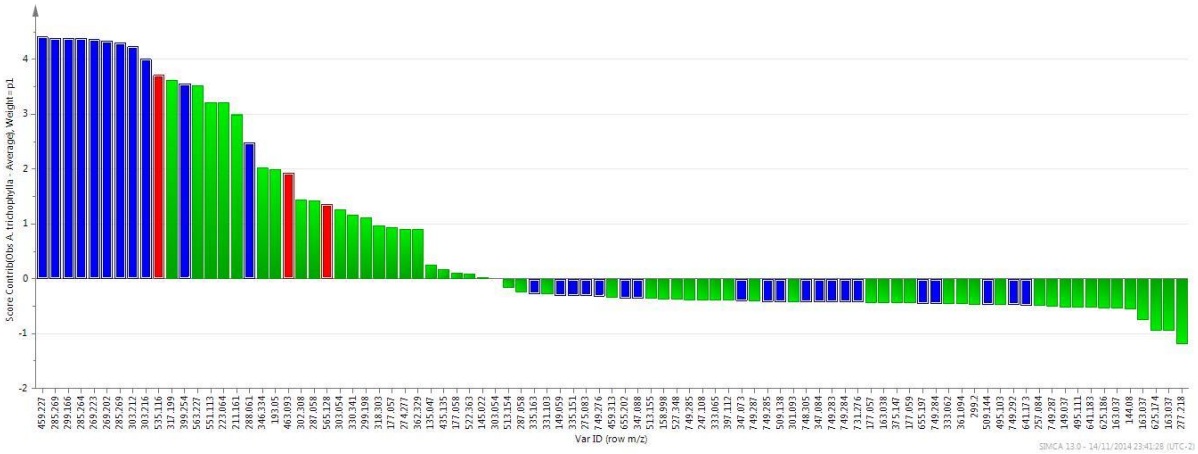


Figura 37 - Variáveis geradas pelo *Contribution Plot* para *A. trichophylla*.

Simultaneamente, podemos gerar o gráfico de variáveis importantes na inibição das enzimas inflamatórias (*Coefficient Plot*) (Figura 38).

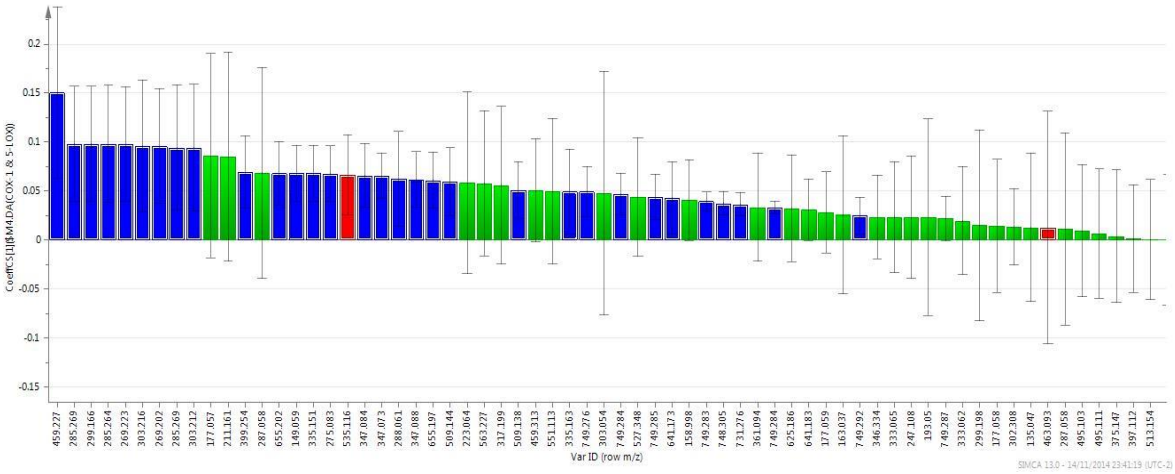


Figura 38 - Variáveis geradas pelo *Coefficient Plot* para inibição de COX-1 e 5-LOX

Aqui, é importante informar que é recomendável que sempre escolhamos as variáveis do *Contribution Plot* que, evidentemente, tenham um valor positivamente correlacionado com a resposta escolhida (eixo Y do gráfico *Contribution Plot*), de preferência que tenham uma média com valor alto e, por fim, que não assumam valores negativamente correlacionados ou com um desvio muito extenso (barras finas). Deste modo, na Figura 38, as barras em azul demonstram as variáveis que satisfazem as condições ditas logo anteriormente. Pode-se notar que muitas destas variáveis não estão presentes em *A. trichophylla* (negativamente correlacionadas no gráfico *Contribution Plot* para *A. trichophylla*). Assim sendo, neste caso, deve-se focar o estudo apenas nas variáveis positivamente correlacionadas em ambos os gráficos, satisfazendo as necessidades deste trabalho. Feito isso, pode-se gerar uma lista detalhada das variáveis que selecionamos, contendo não apenas o valor de *m*/*z* carga de cada pico, bem como o valor de tempo de retenção de cada pico (ou qualquer outra informação que previamente inserimos no SIMCA) e o código de identificação do pico, o qual é gerado pelos *softwares* de processamento de dados, ou em uma planilha de Excel.

Neste momento, podemos notar é que o OPLS-DA fornece muitas variáveis importantes e todas elas podem ser desreplicadas. Evidentemente aquela variável representada por uma barra alta e com baixo desvio (barra fina) é considerada “a mais discriminante” e, supostamente, a mais importante do ponto de vista estatístico para a inibição dupla das enzimas. Entretanto, vale mencionar aqui que nem toda variável importante do ponto de vista estatístico é, necessariamente, importante do ponto de vista bioquímico (Manual do SIMCA 13.0.3). Assim sendo, justifica-se o uso de mais de um modelo para encontrar uma substância química importante, a qual fosse possível de desreplicar e isolar. Além disso, recomenda-se a

utilização de outros *softwares* de outros algoritmos para este intuito. Deste modo, foi utilizado neste também trabalho o *software* Weka, com o qual se fez o *Decision Tree* para os seis melhores modelos descritos anteriormente. Neste estudo, a *Decision Tree* gerada forneceu estatisticamente os valores de amostras classificadas corretamente por meio de uma análise de predição "*10 fold cross-validation leave-one-out*". Ela também fornece valores de RMSEcv. Estes valores podem ser vistos no Quadro 3.

	OPLS-DA	OPLS-DA	Weka	Casos Corretos
Model	Q^2	RMSEcv	RMSEcv	
MetAlign Orbitrap neg + MSClust	**1.00**	**0.02**	**0.34**	**13**
MetAlign Orbitrap pos + MSClust	**0.99**	**0.04**	**0.34**	**13**
Mzmine Orbitrap 3 neg	**0.97**	**0.06**	**0.50**	**11**
Mzmine TOF 4 neg	**0.92**	**0.10**	**0.56**	**10**
Mzmine TOF 1 pos	**0.87**	**0.12**	**0.53**	**7**
Mzmine TOF 4 neg + MSClust	**0.79**	**0.16**	**0.50**	**11**

Quadro 3 – Comparação dos valores de RMSEcv gerados pela *Decision Tree* (Weka) e OPLS-DA para os melhores modelos escolhidos

É esperado que os valores de RMSEcv gerados pela *Decision Tree* sejam piores que do OPLS-DA, visto que aqui não houve deleção de variáveis não importantes (VIP < 1,0). Ainda assim, poderíamos notar que, de maneira geral, os melhores modelos do OPLS-DA também são os melhores do *Decision Tree*. De qualquer modo, a informação mais interessante, e a que interessa neste trabalho, gerado pela *Decision Tree*, é a sua Figura propriamente dita. Assim sendo, para todos os modelos foi gerada uma Figura que mostra qual é a variável que fornece apenas uma resposta classificatória, ou seja, COX-1 e 5-LOX, ou apenas 5-LOX, em uma determinada condição de teste, que neste caso é a área sob o pico do íon detectado. Assim sendo, é verificado se, por exemplo, uma dada variável X ao conter uma área sob o pico ≥ que P, pertencerá à uma dada classe Y. Neste caso, X é a variável, pico codificado de acordo com a tabela gerada pelos processamentos (não mostrada), P é a área sob o pico, e Y é a classificação, que pode ser COX-1 e 5-LOX ou apenas 5-LOX. É importante lembrarmos que as áreas foram normalizas pela amplitude e, desta forma, assumiram valores de 0 a 1. Por fim, neste ponto é de suma importância mencionarmos que a substância mais descriminante encontrada pela *Decision Tree* para o modelo "MZmine TOF 4 neg" é a mesma do modelo "MZmine + MSClust TOF 4 neg" (Figura 39.a e 39.b), evidenciando que o agrupamento gerado pelo MSClust não eliminou uma variável importante, visto que os picos de código 545

e 56 mostraram ser os mesmos nas tabelas de códigos de picos, pois os códigos são diferentes por causa da tabela que o *MSClust* fornece, a qual é menor e, portanto, contém menos picos (ou seja, a numeração é menor). Isso pode ser provado pelo gráfico de *loadings* do OPLS-DA, o qual mostra as variáveis (Figuras 40 e 41) para o modelo "MZmine TOF 4 neg", mostrando que o código 545 pertence a uma substância de *m/z* desprotonada 523.218, a qual inclusive demonstra ser um bom discriminante também no OPLS-DA, visto que encontra-se ao mesmo lado (correlacionada) das amostras que mostraram atividade COX-1 e 5-LOX no gráfico de *scores*. O *loading* de m/z 461.185 desprotonada será explicada mais adiante.

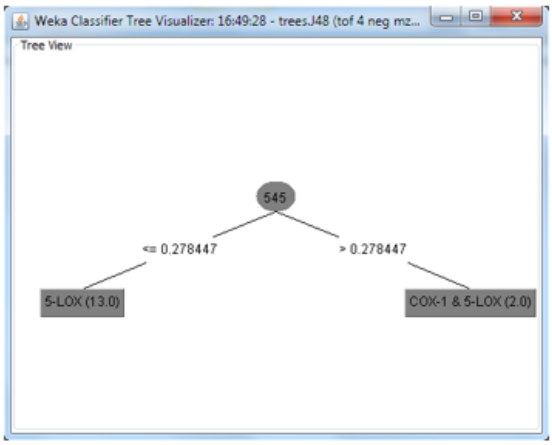


Figura 39.a – substância discriminante para o modelo "MZmine 4 TOF" apontada pela *Decision Tree*

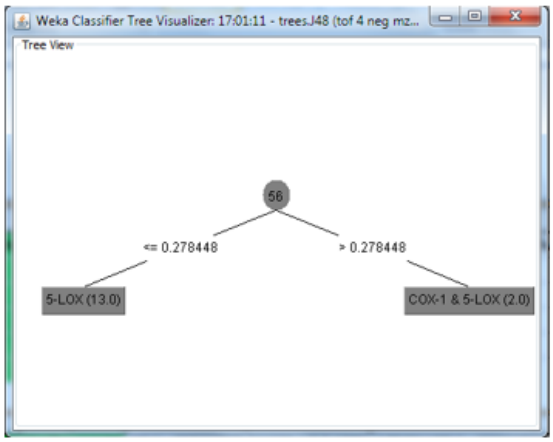


Figura 39.b – substância discriminante para o modelo "MZmine + MSClust TOF 4 neg" apontada pela *Decision Tree*

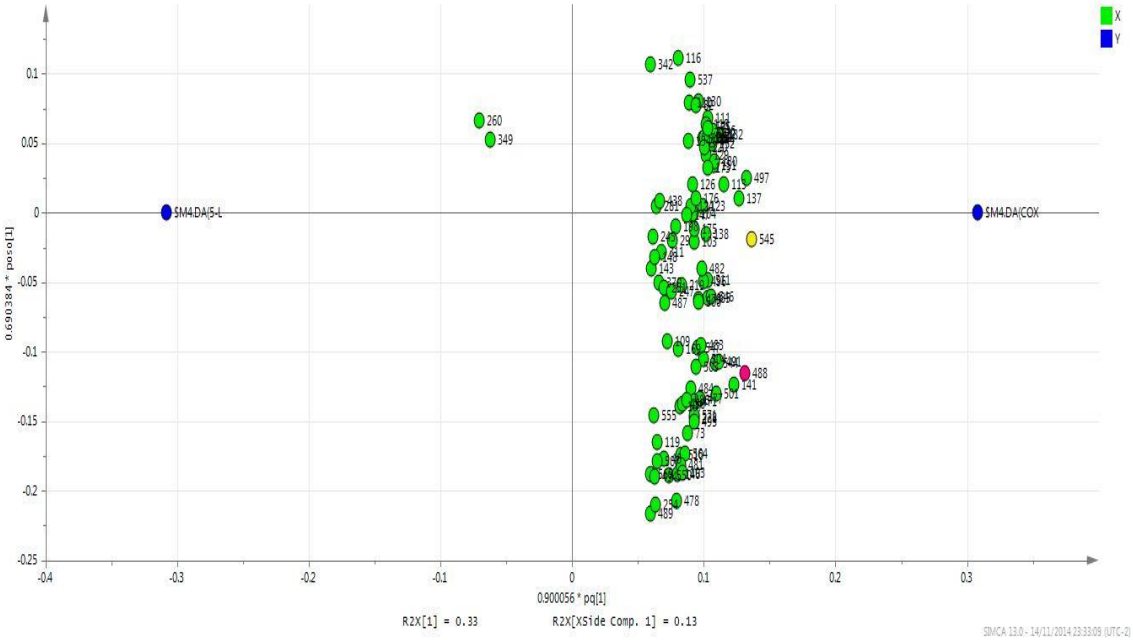


Figura 40 – Gráfico de *loadings* evidenciando a substâncias de código 545 (*m/z* 523) para o modelo "MZmine TOF 4 neg".

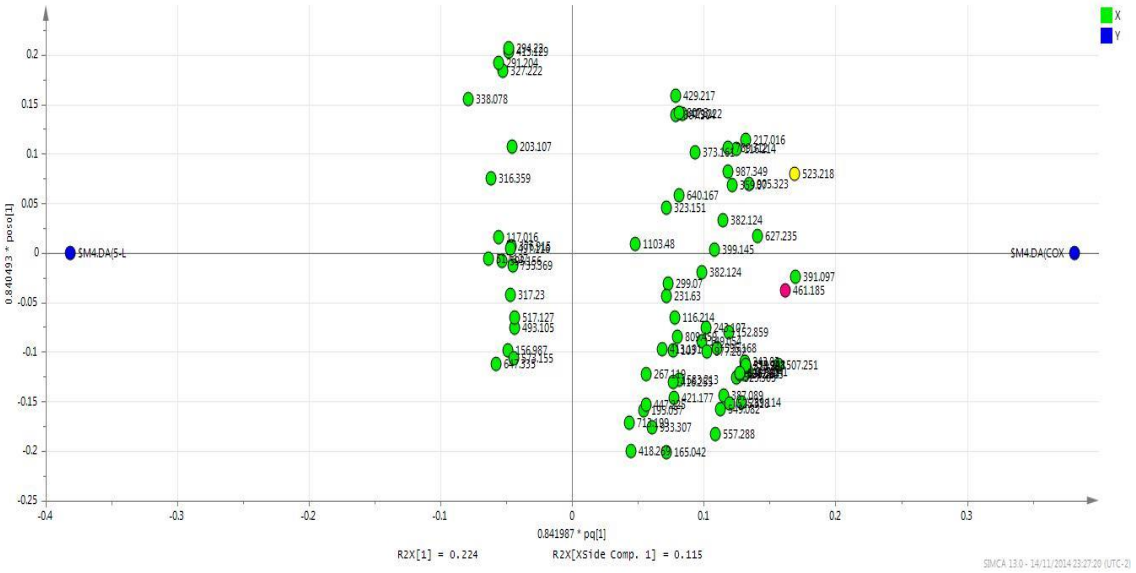


Figura 41 - Gráfico de *loadings* evidenciando a substâncias de *m/z* 523 (código 545) para o modelo "MZmine TOF 4 neg".

O Quadro 4 mostra as variáveis apontadas pela *Decision Tree* como as mais importantes para a classificação COX-1 e 5-LOX (inibição dupla) com base nos dados químicos provenientes dos 6 melhores modelos criados.

Modelo	Código	*m/z*	tr	5 + OPLS-DA ?
MetAlign Orbitrap neg + MSClust	664	219.102	14.12	não
MetAlign Orbitrap pos + MSClust	290	623.210	9.61	sim
Mzmine Orbitrap 3 neg	1398	373.093	8.84	não
Mzmine TOF 4 neg	545	523.218	18.02	não
Mzmine TOF 1 pos	174	459.227	26.41	sim
Mzmine TOF 4 neg + MSClust	56	523.218	18.02	não

Quadro 4 – Variáveis apontadas pela *Decision Tree*

Foram escolhidos tantos picos quanto possível, apontados pela quimiometria, totalizando por volta de aproximadamente 50 variáveis, para se fazer a desreplicação. Esta era realizada observando-se o perfil do espectro no UV do pico, bem como a sua fragmentação. As *m/z* destes picos também foram checadas com o banco de massas de AldamaDB construído de acordo com revisão literária. Por fim, quando se conseguia realizar a desreplicação, a substância era checada se pertencia ao banco de dados *in-house* da equipe científica AsterBioChem, podendo-se comparar o tempo de retenção.

Muitos dos picos apontados são de difícil análise tanto pelo espectro no UV quanto pelo padrão de fragmentação, pois são de baixa intensidade. Por meio do UHPLC-PDA-Orbitrap os íons foram fragmentados via *Higher-energy Collisional Dissociation* (HCD) com energia de coalisão normaliza (*Normalized Collision Energy* – NCE) ajustada em 35%. Pode-se ver em detalhes na Figura 42 extraída da página da internet www.planetorbitrap.com – (*Thermo Fisher*).

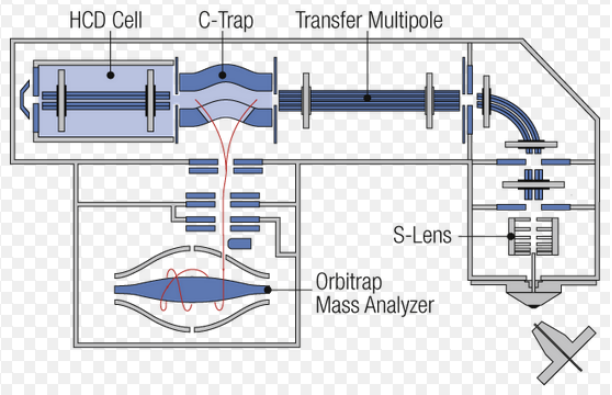


Figura 42 - Esquema de compartimentos do Orbitrap Exactive Plus (Extraído de planetorbitrap.com - *Thermo Fisher*).

Assim sendo, está técnica é extremamente vantajosa em relação à CID, o qual sofre com perdas por *low-mass cut-off*, quando fragmentos muito pequenos não são detectados. Neste trabalho, o *low-mass cut-off* foi de 27%, ou seja, fragmentos com massa inferior a 27% do íon pai não foram detectados. A CID ocorre em uma cela específica e fragmenta o íon após a sua excitação ou ativação colisional com um gás inerte. Entretanto, a cela de HCD por si só não resolve muitos problemas de desreplicação quando há muitos picos eluindo em tempos de retenção muito próximos. Para que ela fosse extremamente útil, necessariamente a resolução cromatográfica teria de ser perfeita, enviando apenas um íon de cada vez ao HCD. Uma resolução cromatográfica depende de muitos parâmetros e não convém abordá-la neste trabalho. De qualquer modo, mesmo com íons eluindo em tempos de retenção próximos, o problema poderia ser contornado por um "filtro" de massas, o qual é possível com um analisador Quadrupolar, o qual seleciona a *m*/*z* de interesse. Infelizmente, o espectrômetro de massas Orbitrap Exactive Plus (*Thermo Scientific*) da AsterBioChem não conta com um Quadruplo acoplado.

Apesar das desvantagens em não se poder contar com um Quadrupolo acoplado ao Orbitrap, com bastante atenção e conhecimento prévio, pode-se fazer um estudo de desreplicação utilizando-se do UHPLC-PDA-Orbitrap, quando o perfil químico não se

encontra demasiado complexo (com muitos íons). Na Figura 43 podemos ver o perfil químico do espectro no UV registrado em 254 nm no *software* Xcalibur™ (*Thermo Scientific*) para o ET da espécie *A. trichophylla*.

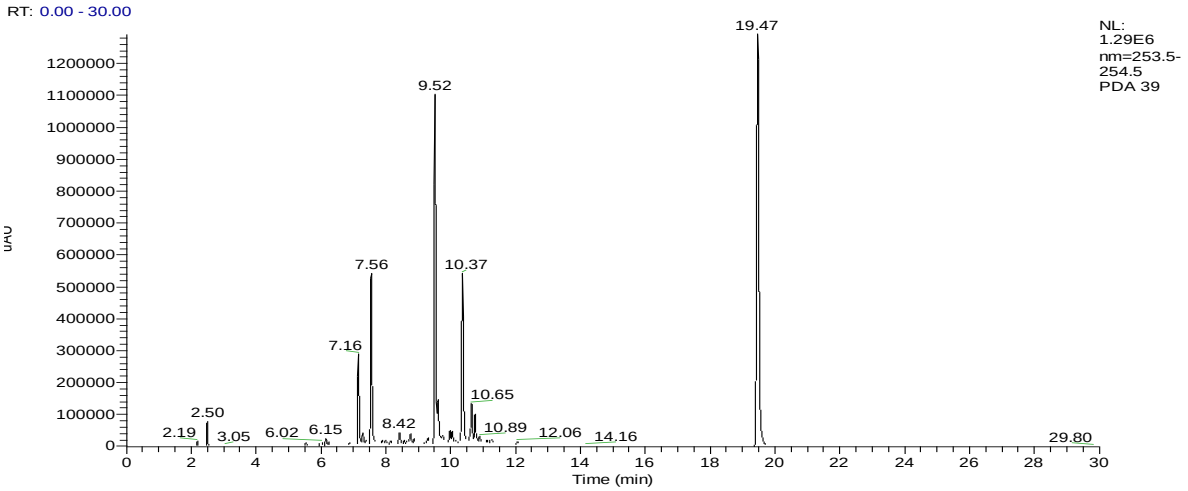


Figura 43 - Perfil químico do espectro de ultravioleta registrado em 254 nm para ET de *A. trichophylla*

Neste momento, é importante mencionar algumas abordagens que foram utilizadas visando-se facilitar a desreplicação e posterior isolamento das substâncias alvo. Primeiro, não apenas o ET foi analisado na desreplicação, mas também as frações resultantes de sua partição, descrito em "Materiais e Métodos". Com isso, os picos apontados pela quimiometria não foram apenas procurados no ET, mas também nas frações AcOEt (fração/partição AcOEt), diclorometano (fração/partição Dicloro), e aquosa (fração/partição aquosa). Segundo, foram acrescentados a estas frações os padrões internos cafeína e flavona, os quais aparecem no cromatogramas da Figura 44 e saem nos tempos 7,56 e 19,47 min, respectivamente. Com tais padrões internos, esperaríamos obter maior precisão na desreplicação quando utilizássemos o tempo de retenção, podendo-se calcular o tempo relativo das substâncias. Na Figura 44 vemos apenas estes padrões em 254 nm no espectro de UV.

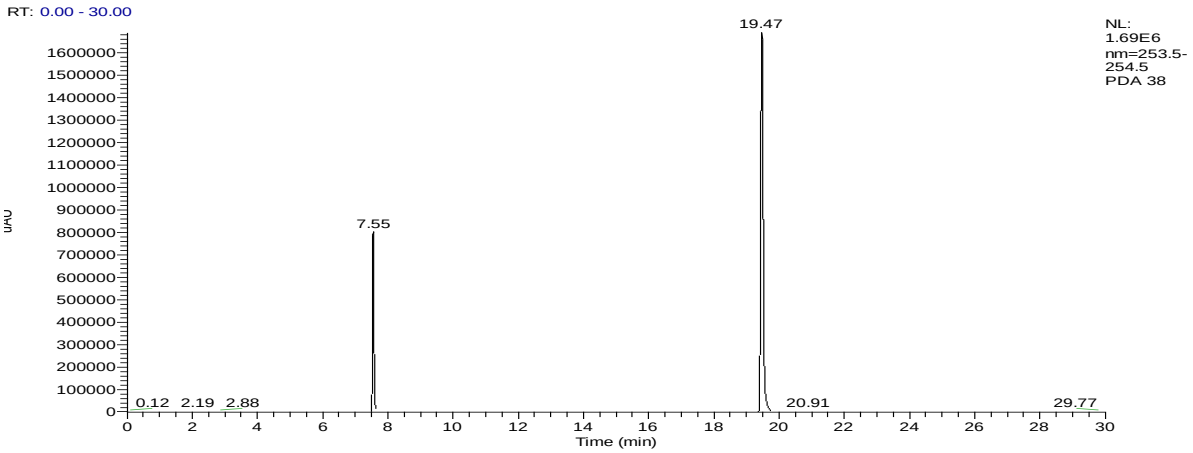


Figura 44 - Perfil químico do espectro de ultravioleta registrado em 254 nm para cafeína e flavona, respectivamente.

Devemos também informar que no UHPLC-PDA-Orbitrap, utilizado neste trabalho, o *delay* (atraso) observado no TIC em relação ao espectro de UV é praticamente insignificante (da ordem de 0,05 centésimos de minuto), apesar de variar de pico para pico ao longo do tempo de corrida. Por exemplo, o *delay* observado para a cafeína foi de 0,03, já para a flavona foi de 0,05. Outra observação que tivemos foi que os perfis químicos registrados em 254 nm Figura 45 são muito semelhantes ao observados em 325 nm no UV Figura 46.

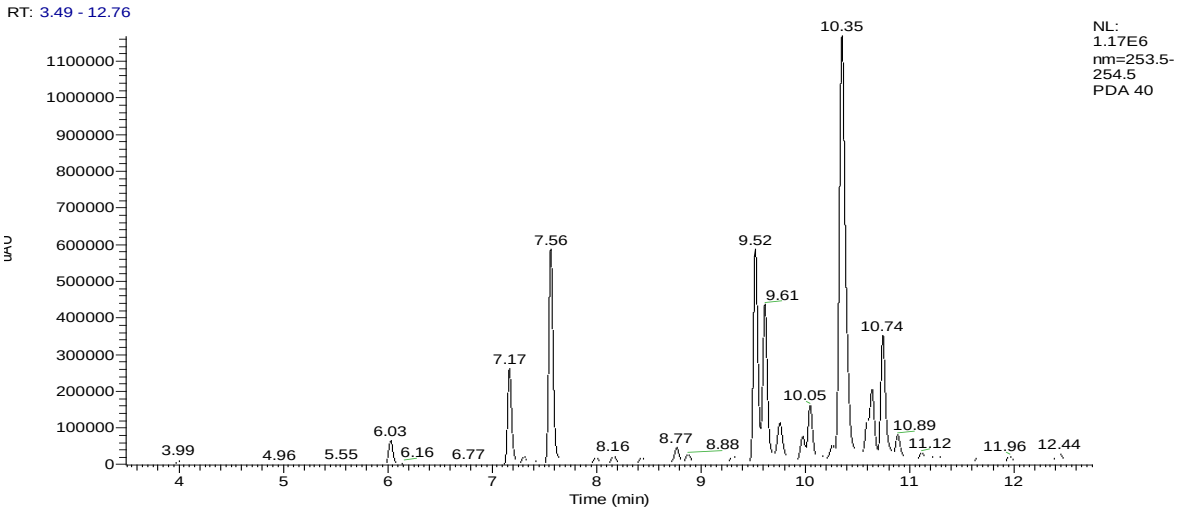


Figura 45 - Perfil químico registrado em 254 nm para *A. trichophylla.*

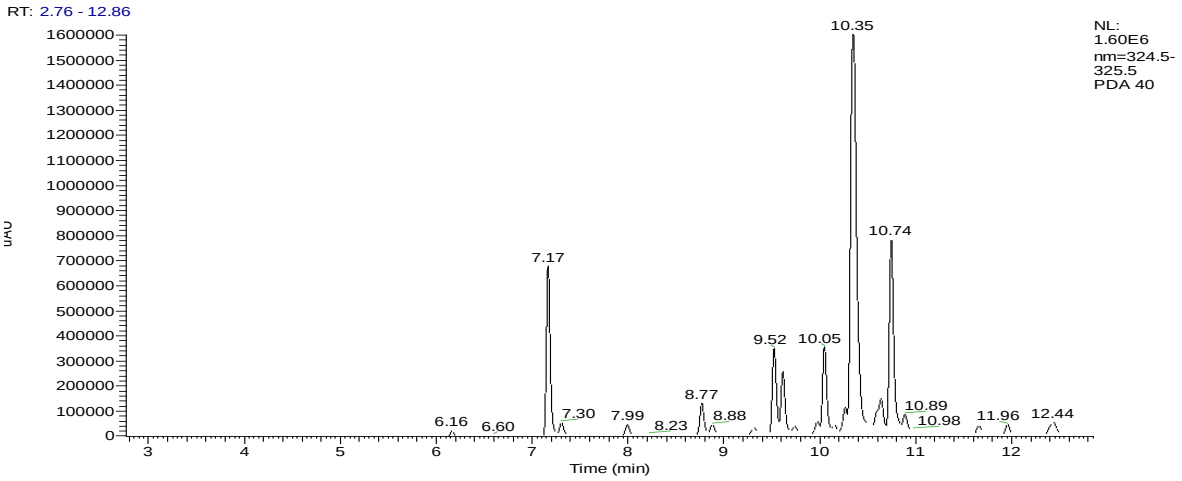


Figura 46 - Perfil químico registrado em 325 nm para *A. trichophylla.*

De qualquer modo, o perfil no UV dos picos foi registrado por meio do DAD que permite registrar múltiplos comprimentos de onda, traçando um perfil. O TIC foi observado nos modos positivo e negativo. Além disso, não apenas os picos apontados pela quimiometria como discriminantes foram desreplicados, mas, fazendo melhor proveito das análises, alguns picos majoritários também foram desreplicados. Assim sendo, por exemplo, o pico que elui em 10,05 min na fração AcOEt e observado na Figura 47 trata-se do ácido 1,4-di-*O*-cafeoilquínico, de acordo com o perfil de UV característico para esta substância (Figura 47), e de acordo com o padrão de fragmentação reportado pela literatura (Clifford *et al*., 2003 e 2005). Deste modo, por exemplo, podemos visualizar o espectro do ácido 1,4-di-*O*-cafeoilquínico no modo positivo (Figura 48), seu espectro de fragmentação gerado pelo HCD no modo positivo (Figura 49), bem como aquele gerado pelo HCD no negativo (Figura 50).

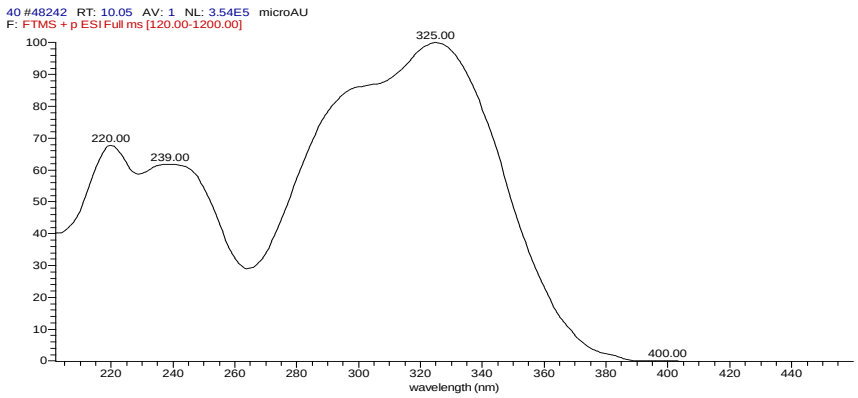


Figura 47 - Perfil de UV típico de ácido cafeoilquínico.

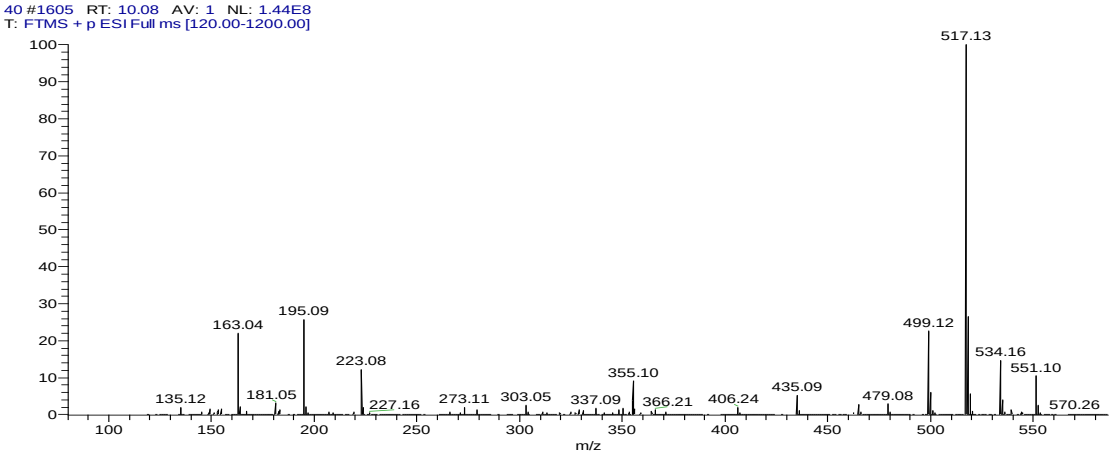


Figura 48 – Espectro de massas do ácido 1,4-di-*O*-cafeoilquínico gerado no modo de ionização positivo.

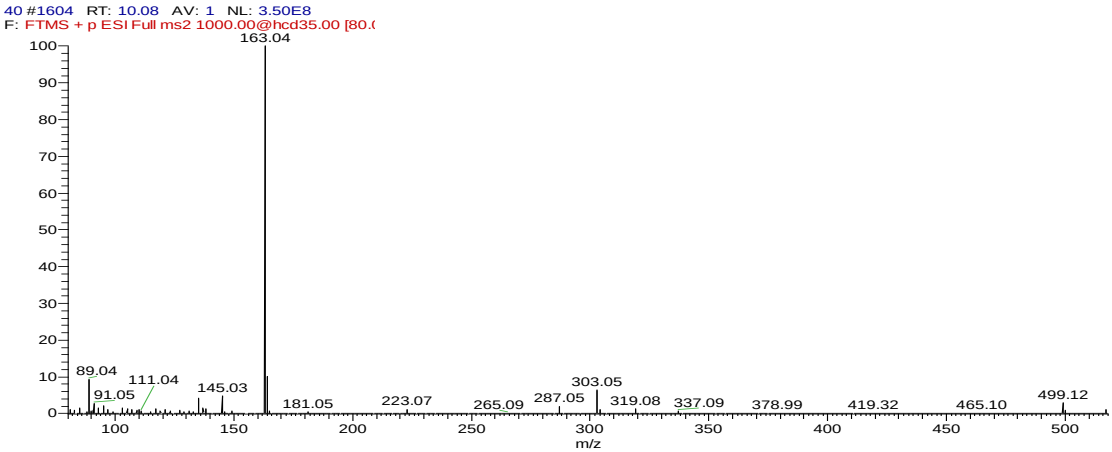


Figura 49 - Espectro de fragmentação do ácido 1,4-di-*O*-cafeoilquínico gerado pelo HCD no modo positivo.

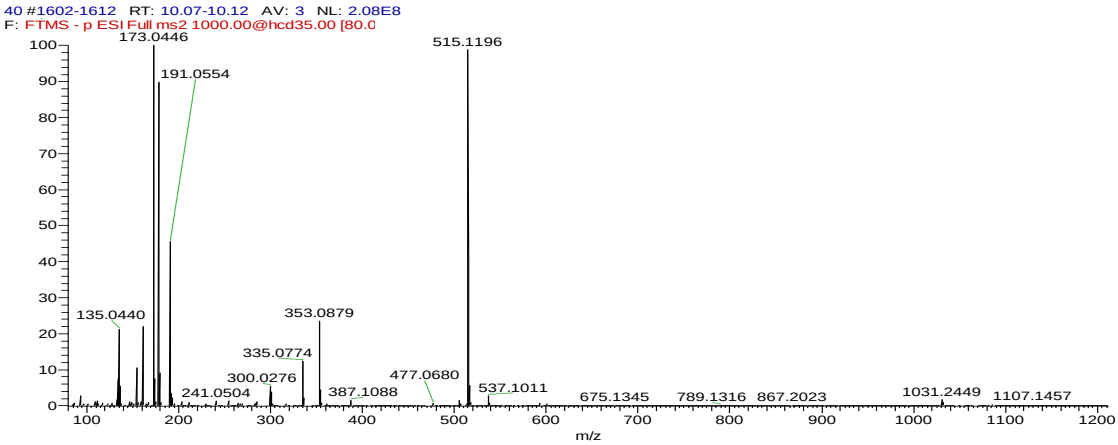


Figura 50 - Espectro de fragmentação do ácido 1,4-di-O-cafeoilquínico gerado pelo HCD no modo negativo

Quanto aos dados analisados provenientes do HPLC-DAD-TOF a fragmentação dos picos apontados pela quimiometria ou dos majoritários foi realizada via HPLC-DAD-Ion-

Trap. Esta foi a alternativa encontrada, uma vez que o analisador TOF não fragmenta massas específicas, ao passo que este Ion-Trap poderia fragmentá-las sem a necessidade de um quadrupolo acoplado. É certo que a mesma estratégia também poderia ter sido utilizada para o UHPLC, desde que se mantivesse a mesma coluna no experimento (coluna ACE – ver Material e Métodos), mas não foi o caso, de modo que, sucintamente, foi realizada uma comparação entre técnicas UHPLC-Orbitrap-HCD e HPLC-Ion-Trap, sendo que ambos analisadores registram o espectro no modo positivo e negativo simultaneamente. Fato é que, para a desreplicação, o TOF e o Orbitrap são excelentes analisadores quando possuem um quadrupolo acoplado a eles, entretanto, o Orbitrap Exactive Plus ainda tem a vantagem de possuir a cela HCD, com energia de fragmentação normalizada. A resolução e acurácia dos analisadores será discutida posteriormente.

Por meio do analisador Ion-Trap podemos selecionar o íon de interesse controlando a amplitude da radiofrequência do instrumento. A física do Ion-Trap é demasiada complexa e uma breve descrição será feita aqui. Esta armadilha de íons é do tipo *Paul Trap*, onde a captura dos íons depende da aplicação de amplitude de radiofrequência entre um eletrodo anelar e dois eletrodos terminais, confinando em um caminho cíclico descrito pela equação de Mathieu o movimento dos íons (Vessecchi, *et al*., 2011). O dispositivo é semelhante a um quadrupolo anelar (3D quádruplo), que funciona por meio de uma corrente contínua estática e uma radiofrequência alternante. De maneira bastante simplificada, na física do Ion-Trap-MS à medida que a voltagem da radiofrequência aumenta, íons de maior *m*/*z* são ejetados e detectados. Wolfgang Paul (laureado com Nobel de física em 1989) e Helmut Steinwedel, com o trabalho "*Ein neues Massenspektrometer ohne Magnetfeld*", descrevem este espectrômetro de massas do tipo *Paul Ion Trap* pela primeira vez em 1953, patenteando-o em 1960. Importante mencionar que o controle da amplitude da radiofrequência permite também selecionarmos íons e fragmentos de interesse para realizarmos estudos de fragmentação de estruturas químicas.

A título de curiosidade, Orbitrap é uma marca registrada pela *Thermo Scientific*, e este espectrômetro, que também é uma armadilha de íons, é do tipo *Kingdon Trap* cuja "frequência das oscilações harmônicas dos íons presos em órbita ao longo do eixo do campo eletrostático é independente da velocidade do íon, porém inversamente proporcional à raiz quadrada da *m*/*z*, de modo que esta armadilha pode ser operada como um espectrômetro de massas, utilizando a detecção da imagem corrente e a transformada de Fourier (*Fourier Transform*) do sinal no domínio tempo", tal como mencionando por Vessecchi e

colaboradores (2011). Um melhor discernimento podemos ter ao observarmos a fórmula que descreve esta armadilha:

$$\omega = \sqrt{\frac{z}{m} \cdot k}$$

Onde ω é a frequência da oscilação harmônica *e k* é a constante instrumental. Assim sendo, a trajetória do íon combina a rotação em torno do eletrodo central axial com as oscilações harmônicas ao longo dele (Figura 51 – Extraída de Hu *et al*., 2005). A frequência ω destas oscilações harmônicas ao longo do eixo z depende apenas da relação massa-carga do íon e da constante instrumental (Hu *et al*., 2005).

Existe também armadilhas de íons do tipo *Penning Trap*, os quais são costumeiramente chamados de analisadores de massa por Ressonância Ciclotrônica (*Fourier Transform Ion Cyclotron Ressonance Mass Spectrometry*) onde os íons são presos em um movimento circulatório, dentro de um campo magnético fixo, e cuja relação massa carga (*m*/*z*) pode ser determinada. Não convém descrevê-lo aqui em maiores detalhes, pois este dispositivo não fora utilizado neste trabalho.

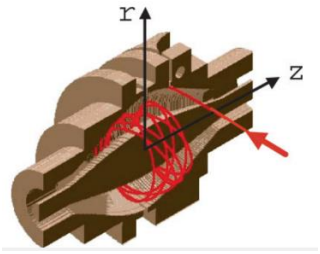


Figura 51 - Esquema de movimento dos íons no Orbitrap (Extraída de Hu *et al*., 2005).

Sendo assim, voltando ao experimento com o Ion-Trap, por meio dele pode-se realizar um estudo de MRM, tal como descrito em "Material e Métodos", onde foram gerado íons fragmentos até a terceira geração (MS^3). Para isso, primeiramente é realizado um MS^2 onde os fragmentos de todas as moléculas protonadas ou desprotonadas são registradas. Após isso, podemos selecionar o fragmento de interesse e realizar um estudo direcionado para ele, fragmentando-o em MS^3. Deste modo, por meio da HPLC-DAD-Ion-Trap fez-se a desreplicação, registrando o perfil de UV dos picos apontados pela literatura e de alguns picos majoritários. A Figura 52 mostra o perfil químico da fração AcOEt registrado em 254 nm para

A. trichophylla. Em 10,4 e 43,3 min eluem os padrões internos cafeína e flavona, respectivamente.

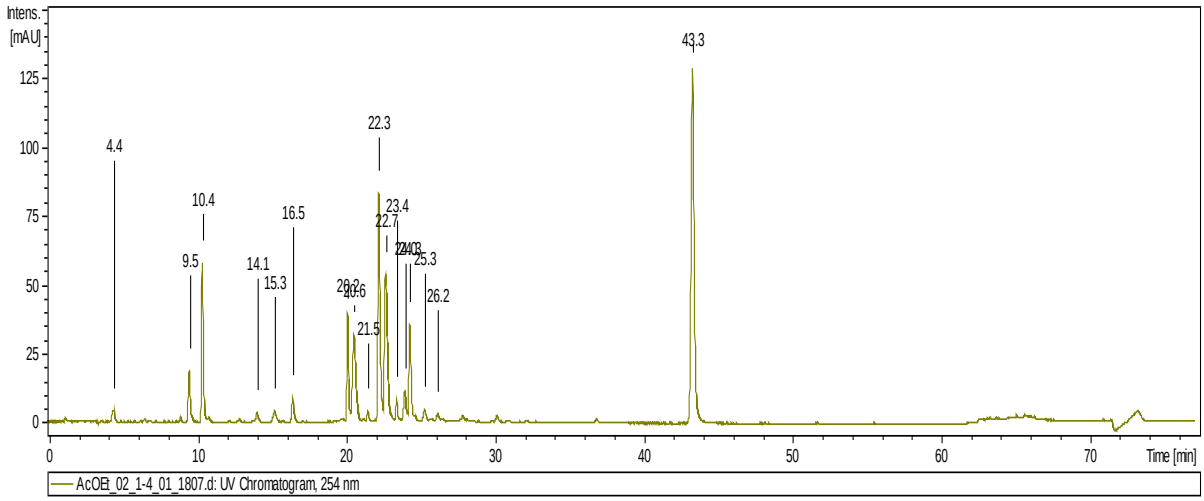


Figura 52 - Perfil químico da fração AcOEt registrado em 254 nm para *A. trichophylla no* Ion-Trap.

Neste ponto já podemos informar que percebemos também que a fração mais promissora para fazermos a desreplicação foi a fração AcOEt. Na fração diclorometano não observamos sequer picos majoritários tanto pelo perfil de UV quanto pelo TIC, donde supomos que ela é rica em di e triterpenos, os quais tem dificuldade de se ionizar em fontes ESI, além de terem baixa absorção no UV (abaixo de 220 nm), e, portanto, podem não ser observadas pela absorção do ácido acético ou fórmico colocado na fase móvel. Uma alternativa para este problema seria a aquisição de uma fonte de ionização por APCI (*Atmospheric Pressure Chemical Ionization*) a qual possibilitaria uma melhora na detecção de metabólitos de média a baixa polaridade (Zhou *et al*., 2012).

Independentemente, mesmo com a fonte ESI operando em todas as nossas análises foi possível satisfatoriamente realizarmos um estudo de desreplicação que satisfizesse aos propósitos deste trabalho. Assim sendo, por exemplo, o pequeno pico que aparece no tempo 22,7 min no perfil químico da Figura 52, foi apontado pela quimiometria como discriminante para atividade COX-1 e 5-LOX, assim como os picos que saem em 25,3 e 26,2 min, e ele se trata do kaempferol-3-*O*-glucuronideo, um flavonol glicosilado, de acordo com o seu perfil de UV (Figura 53) e o padrão de fragmentação (Figura 54, MS^2, e 55, MS^3) apresentado e confirmado na literatura (Tsimogianins *et al*., 2007; Sakakibara *et al*., 2003; Markham, 1982). Essa substância também foi apontada no modelo “MZmine TOF 4 neg” e aparece em

destaque (cor rosa) no gráfico de *loadings* da Figura 41, gerado com dados do modo negativo (*m*/*z* 461).

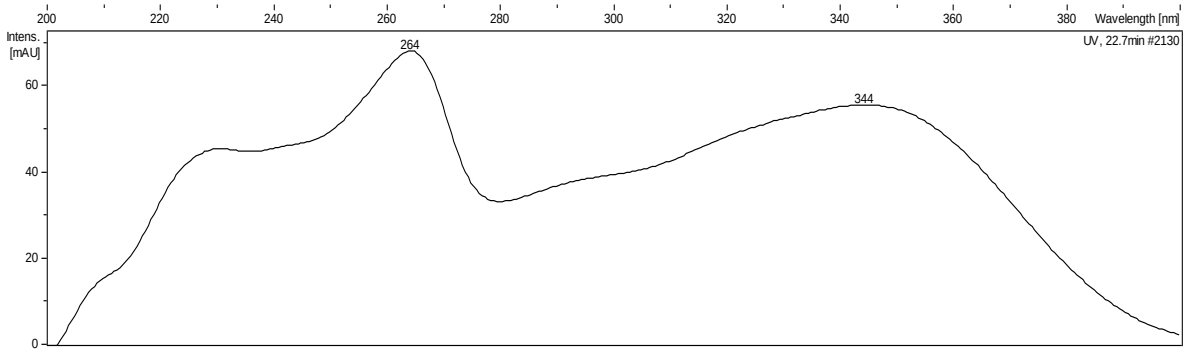


Figura 53 – Perfil de UV para kaempferol-3-*O*-glucuronideo.

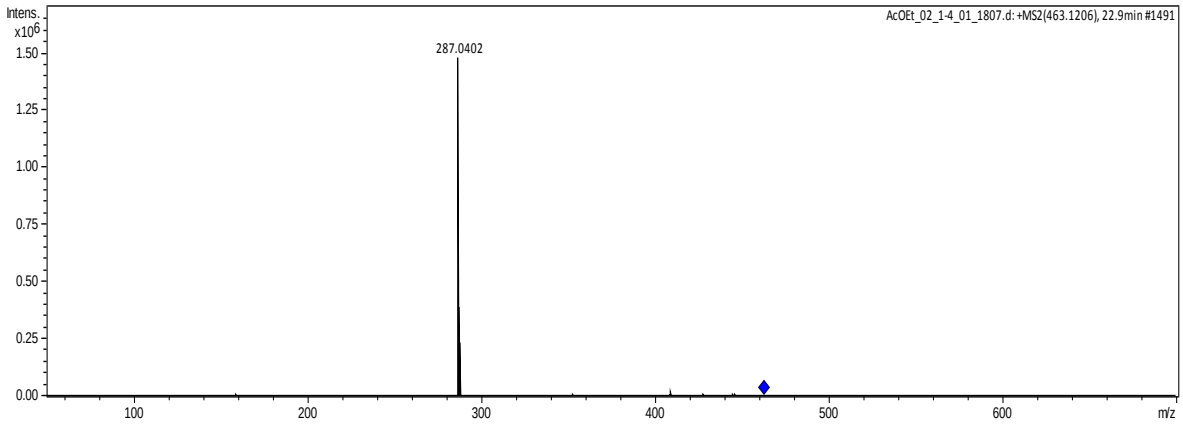


Figura 54 – Espectro MS² do kaempferol-3-*O*-glucuronideo, evidenciando o fragmento 287 gerado a partir da molécula protonada de *m/z* 463.

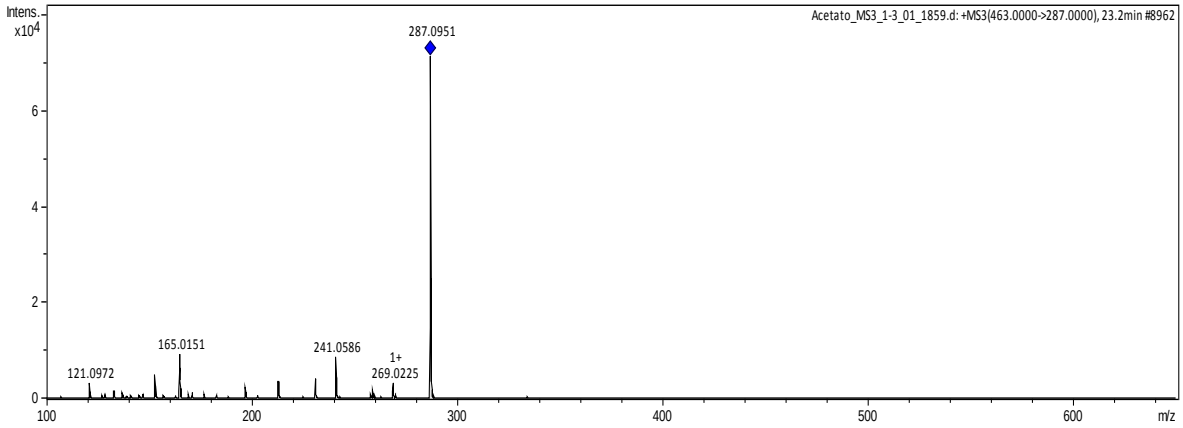


Figura 55 - Espectro MS³ do kaempferol-3-O-glucuronideo, evidenciando fragmentos que se formam a partir do íon de *m/z* 287, o qual foi gerado a partir da molécula protonada de *m/z* 463.

Portanto, utilizando-se de perfil de UV e padrões de fragmentação registrados no modo positivo e negativo tanto para o Orbitrap quanto Ion-Trap foi possível desreplicar 12 substâncias químicas, das quais três são apontadas pela quimiometria como possíveis inibidoras de COX-1 e 5-LOX. Destas três substâncias, duas foram detectadas e desreplicadas por meio dos dois analisadores. A Tabela 3 sumariza estas substâncias apontadas pela quimiometria com seus respectivos cálculos de acurácia. A Tabela 4 mostra todas as substâncias e os mecanismos de fragmentação nos dois modos de ionização. Em ambas a numeração em parênteses diz respeito ao fato da estrutura pertencer ao banco AldamaDB construído com informações de massas de moléculas protonadas e desprotonadas, fragmentos e adutos de substâncias de *Aldama* da literatura.

Os ácidos cafeoil e dicafeoilquínicos foram desreplicados com base no característico perfil de absorção no UV já descrito anteriormente e com base nos padrões de fragmentação das estruturas químicas, os quais estão muito bem documentados nos trabalhos de Clifford e colaboradores (2003 e 2005). Por exemplo, o primeiro ácido clorogênico que aparece em 9,5 min no HPLC-Ion-Trap foi desreplicadoscomo sendo o ácido 1-*O*-cafeoilquínico, uma vez que esta substância não apresenta fragmento 135 no espectro de MS^2 adquirido no modo negativo, sendo que os ácidos 3, 4 e 5-cafeoilquínicos apresentam. Do mesmo modo, o segundo ácido clorogênico foi desreplicado como sendo o ácido 4-cafeoilquínico, pois este apresenta pico base de *m*/*z* 173 (Clifford *et al*., 2005). Não obstante, neste ponto é de suma importância mencionarmos que estes padrões de fragmentação da literatura foram comparados com os padrões de fragmentação do Ion-Trap, com os quais foram realizados os estudos de RMN. Os dados de fragmentação obtidos com o Orbitrap foram úteis mais para confirmar a desreplicação do Ion-Trap, uma vez que o Orbitrap Exactive Plus utilizado neste experimento não possui um quadrupolo acoplado para selecionar íons de *m*/*z* específicos. Além disso, o analisador de massas utilizado por Clifford e seus colaboradores também é um Ion-Trap (LCQ Deca XP Plus – *Thermo Finnigan*).

Deste modo, dando continuidade à desreplicação dos ácidos dicafeolquínicos, a substância que elui em 21,5 min foi desreplicada como sendo o ácido 1,4-*O*-dicafeoilquínico, uma vez que o seu MS^3 apresenta o pico base de *m*/*z* 173, embora o ácido 4,5 também apresentasse o mesmo fragmento em MS^3, no espectro MS^2 esta substância não apresenta o fragmento de *m*/*z* 335, o qual fora encontrado aqui. Ainda assim, poderíamos ficar em dúvida também em relação ao ácido 3,4-*O*-dicafeoilquínico. Entretanto, esta substância apresenta um tempo de retenção acima de 25 min registrado em nossos dados de tempo de retenção do nosso banco de padrões *in-house*, os quais foram realizados nas mesmas condições

experimentais deste trabalho. A substância que elui em 22,3 min no LC-Ion-Trap foi desreplicada como sendo o ácido 3,5-*O*-dicafeoilquínico, visto que o padrão de fragmentação por ela apresentada no espectro MS^3 é muito semelhante ao informado pela literatura (Clifford *et al*., 2005). A presença do fragmento de *m*/*z* 173 em MS^3 foi decisiva para esta desreplicação. A despeito disso, um detalhe interessante é que o pico base apresentado pela fragmentação via Orbitrap para esta estrutura é de *m*/*z* 191, o mesmo pico base do espectro MS^3 do Ion-Trap, mas diferente do espectro MS^2, embora ainda contenha o fragmento característico de dicafeoilquínicos *m*/*z* 353. Além disso, no modo de ionização positivo, podemos observar que todos os fragmentos pico base do Orbitrap para os ácidos cafeoilquínicos é 163, diferentemente do observado no espectro MS^3 do Ion-Trap. Também podemos notar que o HCD parece gerar uma maior quantidade de fragmentos em um mesmo espectro, com fragmentos de menor *m*/*z*. Isto ocorre não somente por causa do *low-mass cut-off* do Ion-Trap, mas também por causa que a energia de coalisão normalizada da HCD nada mais é que uma energia de coalisão que proporcional à *m*/*z* do íon. Deste modo, íons com maior massa tendem a fragmentar mais que íons com menor massa, além de não ocorrer perda por *low-mass cut-off*, o que poderia levar o estudo a estágios de MS^n mais avançados, caso este Orbitrap contasse com um quadrupolo acoplado. Por fim, o ultimo ácido cafeoilquínico da Tabela 4 só poderia ser o ácido 3,4-*O*-dicafeoilquínico, o qual apresenta o fragmento de *m*/*z* 335, de acordo com o espectro MS^2 de Clifford e colaboradores (2005) e um maior tempo de retenção, de acordo com nossas condições analíticas e banco de dados. Ainda assim, há de considerarmos que o MS3 que apresentamos não está de acordo com a literatura, onde observamos um pico base *m*/*z* 191, diferentemente de 173, apesar disso, o HCD do Orbitrap apresenta um pico base *m*/*z* 173, podendo evidenciar um caso de complementariedade e auxilio na desreplicação via Ion-Trap.

Tabela 3 – Prováveis substâncias (nomes em inglês) inibidoras de COX-1 e 5-LOX desreplicadas em alguns dos melhores modelos de predição criados via OPLS-DA.

Desreplicação das substâncias apontadas pela quimiometria									
Modelos e substâncias provenientes do TOF-MS (Fragmentação por Ion-Trap-MS/MS)									
Modelo	**RT**	**m/z exp.**	**MS^2 (FA 0.6)**	**MS^3 (FA 0.5)**	**UV**	**classe**	**possível substância**	**m/z teórica**	**Acurácia**
MzMine 1 positive	22.73	463.0902	463 → 287 (bp); 240; 215; 171; 103	287 → 269 [-H_2O], 241 [-H_2O-CO], 231 [-2CO], 213 [-H_2O-2CO], 165 [$^{0,2}A^+$], 153 [$^{1,3}A^+$], 121 [$^{0,2}B^+$]	232sh, 264; 295sh, 345	flavonol	kaempferol-3-O-glucuronide	463.0876	6E-06
MzMine 4 negative ou MSClust 4 negative	22.83	461.0746	461 → 447; 285 (bp); 174.90	447 → 327	232sh, 264; 295sh, 345	flavonol	kaempferol-3-O-glucuronide	461.0731	3E-06
MzMine 1 positive	25.05	535.1113	535 → 329; 287 (bp)	329 [$^{0,2}X_0$] → 311 [-H_2O], 283 [-CO], 255, 227, 121	265; 295sh, 340	flavonol	kaempferol-3-O-(6''-malonyl-glucoside)	535.1088	5E-06
Modelos e substâncias provenientes do Orbitrap-MS (Fragmentação por HCD)									
Modelo	**RT**	**m/z exp.**	**MS-MS via HCD (NCE 35%)**		**UV**	**classe**	**possível substância**	**m/z teórica**	**Acurácia**
MzMine 3 negative	10.92	461.0724	461 → 285		265; 300sh, 333	flavonol	kaempferol-3-O-glucuronide	461.0731	-2E-06
MzMine 3 negative	11.18	491.0833	491 → 315 (bp), 300 [-CH_2]		254, 265sh; 300sh, 344	flavonol	quercetin-3-O-methyl-7-glucuronide (269)	491.0837	-8E-07

Tabela 4 – Todas as substâncias desreplicadas (nomes em inglês) e seus respectivos padrões de fragmentação encontrado nos modos de ionização positivo e negativo via Ion-Trap e Orbitrap (em destaque as prováveis substâncias ativas contra COX-1 e 5-LOX)

Fragmentação via HPLC-Ion-Trap - modo de ionização positivo				
HPLC-Ion-Trap	***m/z exp. Ion-Trap***	***MS^2 fragmentation amplitude 0.6***	***MS^3 fragmentation amplitude 0.5***	***possível substância***
9.5	355.123	355 --> 163 (bp); 117	163 --> 145, 135 (bp), 117	1-caffeoylquinic acid
10.9	354.410	355 --> 330, 219, 163 (bp)	*Não registrado*	4-caffeoylquinic acid
20.2	465.296	465 --> 303 (bp); 257	303 --> 285 [-H_2O], 229 [-2CO], 137 [$^{0,2}A^+$]	isoquercetrin
20.6	479.142	479 --> 303 (bp); 285; 275	303 --> 285 [-H_2O], 275, 257 [-H_2O-CO], 247 [-2CO], 229 [-H_2O-2CO], 165 [$^{0,2}A^+$], 153 [$^{1,3}A^+$], 137	quercetrin-3-glucuronide (271)
21.5	499.386	517 --> 499, 319, 303 (bp), 163	499 --> 337, 319 (bp), 163, 145	1,4-dicaffeoylquinic acid
22.4	**595.182**	**595 (?) --> 551 --> 345; 303 (bp)**	**----------**	***Não desreplicado***
22.3	499.407	517 --> 499, 319, 163 (bp), 145	499 --> 337, 319 (bp), 163, 145	3,5-dicaffeoylquinic acid
22.7	**463.169**	**463 --> 287 (bp)**	**287 --> 269 [-H_2O], 241 [-H_2O-CO], 231 [-2CO], 213 [-H_2O-2CO], 165 [$^{0,2}A^+$], 153 [$^{1,3}A^+$], 121 [$^{0,2}B^+$]**	**kaempferol-3-O-glucuronide**
23.4	479.020	479 --> 317 (bp)	317 [Y_0]--> 302 [-CH_3], 285 [-H_2O], 271, 261, 165 [$^{0,2}A^+$], 153 [$^{1,3}A^+$], 139	isorhamnetin-3-O-glucoside
24.0	**493.098**	**493 --> 317 (bp), 287**	**317 [Y_0] --> 302 [-CH_3], 285 [-H_2O]**	**quercetin-3-O-methyl-7-glucuronide (269)**
24.3	517.419	517 --> 499, 337, 319 (bp), 162, 145	499 --> 499, 337, 319 (bp), 162, 145	3,4-dicaffeoylquinic acid
25.3	**535.340**	**535 --> 329; 287 (bp)**	**329 [$^{0,2}X_0$] --> 311 [-H_2O], 283, 265, 255, 227, 121**	**kaempferol-3-O-(6"-malonyl-glucoside)**
26.2	565.401	565 --> 359; 317 (bp)	359 [$^{0,2}X_0$] --> 341 [-H_2O], 315, 285, 217	tamarixetin-3-O-β-(6"-malonyl-glucopyranoside)
Fragmentação via HPLC-Ion-Trap - modo de ionização negativo				
HPLC-Ion-Trap	***m/z exp. Ion-Trap***	***MS^2 fragmentation amplitude 0.6***	***MS^3 fragmentation amplitude 0.5***	***possível substância***
9.5	353.062	353 --> 191 (bp), 179 (2.54%)	191 --> 173 (bp), 153, 127, 111	1-caffeoylquinic acid

Continua na próxima página

10.9	352.940	353 --> 191, 179 (22.57%), 173 (bp), 135 (4.25%)	191 --> 176	4-caffeoylquinic acid
20.2	463.030	463 --> 301	*Não fragmentou*	isoquercetrin
20.6	477.017	477 --> 301	*Não registrado*	quercetrin-3-glucuronide (271)
21.5	515.173	515 --> 471, 353 (bp), 335 (8.05%), 191 (14.78%)	353 --> 191 (27.76%), 179, 173 (bp), 135	1,4-dicaffeoylquinic acid
22.4	'----------	----------	----------	***Não desreplicado***
22.3	515.403	515 --> 353 (bp), 335 (3.67%), 191 (6.61%)	353 --> 191 (bp), 179 (47.01%), 173, 135	3,5-dicaffeoylquinic acid
22.7	**461.289**	**461 --> 285 (bp)**	***Não registrado***	**kaempferol-3-O-glucuronide**
23.4	477.121	477 --> 357, 314 (bp), 285, 218, 151	357 --> 342, 329, 314, 285, 193 165	isorhamnetin-3-O-glucoside
24.0	**491.345**	**491 --> 462, 315 (bp), 299**	***Não registrado***	**quercetin-3-O-metil-7-glucuronide (269)**
24.3	515.170	515 --> 353 (bp), 335 (7.10%) 317 (5.69%), 173 (13.73%)	353 --> 191 (bp), 179 (19.25%), 173 (36.15%)	3,4-dicaffeoylquinic acid
25.3	**533.135**	**533 --> 489**	**489 --> 285**	**kaempferol-3-O-(6"-malonyl-glucoside)**
26.2	*Não ionizou*	----------	----------	tamarixetin-3-O-β-(6"-malonyl-glucopyranoside)

Continua na próxima página

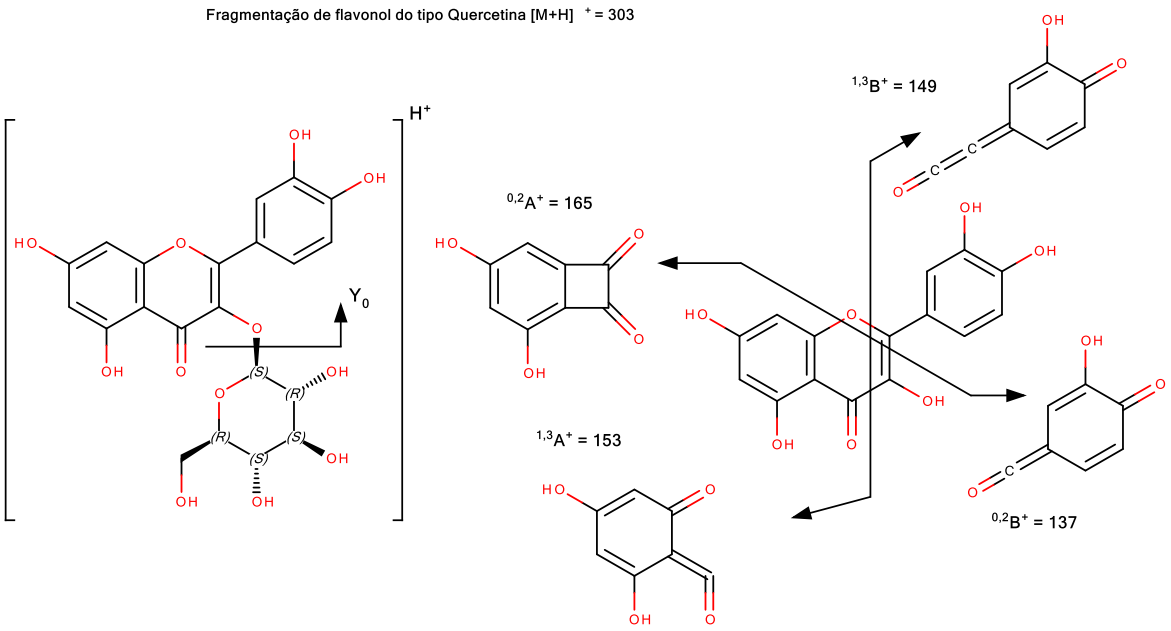
Fragmentação via HCD-Orbitrap - modo de ionização positivo			
UHPLC-Orbitrap	***m/z exp. Orbitrap***	***MS-MS normalized collision energy 35%***	***possível substância***
6.15	355.103	355 --> 163 (bp), 149, 145, 137, 135	1-caffeoylquinic acid
7.17	355.102	355 --> 163 (bp), 145, 135, 117, 89	4-caffeoylquinic acid
9.61	465.103	465 --> 303, 85	isoquercetrin
9.52	479.082	479 --> 303, 85	quercetrin-3-glucuronide (271)
10.05	517.135	499 --> 303, 163 (bp), 145, 89	1,4-dicaffeoylquinic acid
9.98	**595.167**	**595 --> 551, 303 (bp), 287, 223, 163, 127, 91, 85**	***Não desreplicado***
10.26	517.135	517 --> 499, 163 (bp)	3,5-dicaffeoylquinic acid
10.35	**463.087**	**463 --> 287 (bp), 163, 113, 85**	**kaempferol-3-O-glucuronide**
10.60	479.119	479 --> 317 (bp), 302, 287, 163, 151, 127, 97, 85	isorhamnetin-3-O-glucoside
10.64	**493.098**	**493 --> 317 (bp), 302, 287, 163**	**quercetin-3-O-methyl-7-glucuronide (269)**
10.74	517.135	517 --> 499, 317, 287, 163 (bp) 145	3,4-dicaffeoylquinic acid
10.89	**535.109**	**535 --> 287 (bp), 163, 127**	**kaempferol-3-O-(6"-malonyl-glucoside)**
11.12	565.120	565 --> 409, 359, 339, 317 (bp), 287, 223, 177, 163, 159, 127	tamarixetin-3-O-β-(6"-malonyl-glucopyranoside)
Fragmentação via HCD-Orbitrap - modo de ionização negativo			
UHPLC-Orbitrap	***m/z exp. Orbitrap***	***MS-MS normalized collision energy 35%***	***possível substância***
6.15	353.088	353 --> 191 (bp), 179, 161, 135	1-caffeoylquinic acid
7.17	353.088	353 --> 191 (bp), 85	4-caffeoylquinic acid
9.61	463.088	463 --> 300 (bp), 271, 179, 151	isoquercetrin
9.52	477.067	477 --> 301 (bp), 273, 179, 151	quercetrin-3-glucuronide (271)
10.05	515.120	515 --> 353, 335, 300, 191, 179, 173 (bp), 161, 155, 135	1,4-dicaffeoylquinic acid
9.98	**593.152**	**593 --> 550, 505, 433, 387, 300 (bp) 281, 193, 134**	***Não desreplicado***

Continua na próxima página

10.25	515.120	515 --> 353, 191 (bp), 179, 161	3,5-dicaffeoylquinic acid
10.35	**461.073**	**461 --> 285 (bp)**	**kaempferol-3-O-glucuronide**
10.60	477.104	*incompreensível*	isorhamnetin-3-O-glucoside
10.64	**491.083**	**491 --> 315 (bp)**	**quercetin-3-O-metil-7-glucuronide (269)**
10.74	515.119	515 --> 353, 179, 173 (bp) 161, 135	3,4-dicaffeoylquinic acid
10.89	**533.094**	**533 --> 284 (bp)**	**kaempferol-3-O-(6"-malonyl-glucoside)**
11.12	563.105	563 --> 315 (bp)	tamarixetin-3-O-β-(6"-malonyl-glucopyranoside)

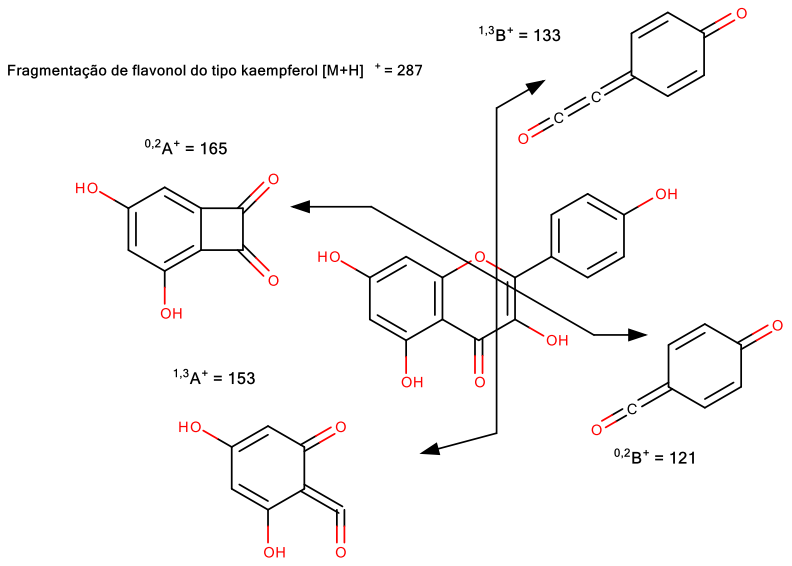
Os padrões de fragmentação descritos na Tabela 3 e 4 podem ser melhor compreendidos através da elucidação de alguns mecanismos de fragmentação descritos logo abaixo, com os respectivos dados de literatura. Além disso, constam também os dados da literatura do espectro no UV:

- Isoquercetrina – Literatura de UV: Sakakibara *et al*., 2003; Literatura MS-MS: Tsimogiannis *et al*., 2007 e Vukics e Guttman, 2010:



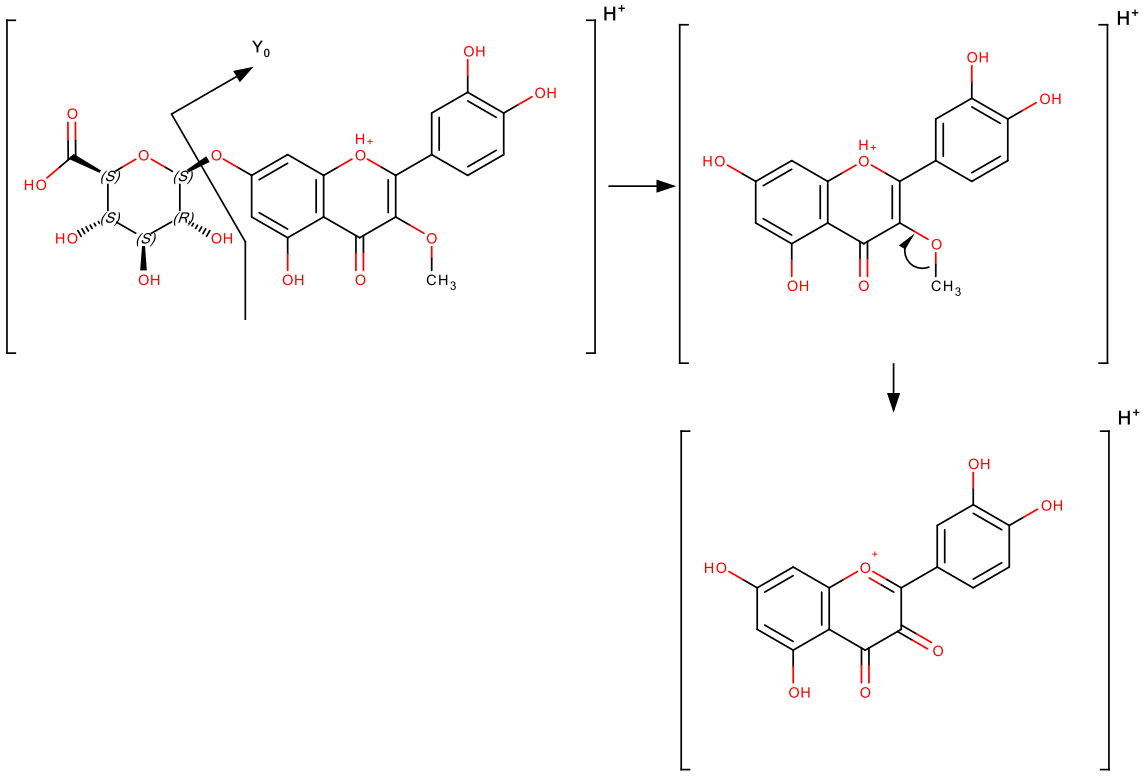
Mecanismo 1 - Fragmentação de flanonol do tipo quercetina. Fragmentos observados para isoquercetina.

Basicamente todos os flavonóis seguem o padrão de fragmentação do mecanismo 1 descrito acima, e todos os fragmentos apresentados nas Tabelas 3 e 4 se correspondem com a literatura. A diferença é sutil e, neste trabalho, a massa do íon precursor foi importante para distinguir entre os diferentes tipos de flavonóis, tais como quercetina, kaempferol, fisetina e flavonol. Um detalhe interessante é que o kaempferol, por conter uma hidroxila a menos que a quercetina, origina os fragmentos de *m*/*z* 133 e 121, para os mecanismos $^{1,3}B^{+}$ e $^{0,2}B^{+}$, respectivamente (mecanismo 2). Os mecanismos de fragmentação dos glicosídeos, bem como as nomenclaturas, estão de acordo com a literatura (Domon e Costello, 1988; Kachilicki *et al*., 2008; Vukics e Guttman, 2008).

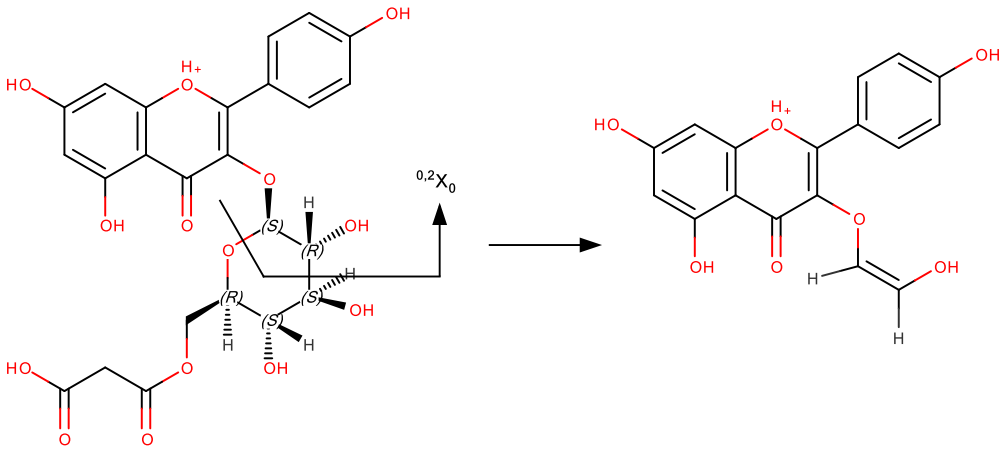


Mecanismo 2 - Fragmentação para flavonol do tipo kaempferol.

Alguns dos mecanismos de fragmentação dos glicosídeos podem ser visualizados nos mecanismos 3 e 4, com destaque para o mecanismo de fragmentação da isorhammetima-3-*O*-glicosideo e kaempferol-3-*O*-(6''-malonil-glucosideo).



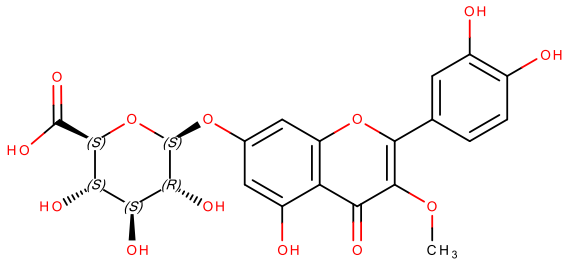
Mecanismo 3 - Fragmentação da isorhammetima-3-*O*-glicosideo.



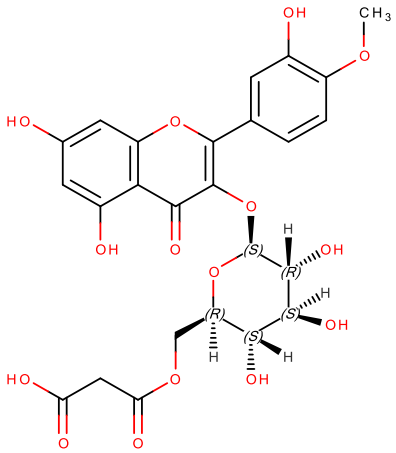
Mecanismo 4 - kaempferol-3-*O*-(6"-malonil-glucosideo).

Abaixo pode-se visualizar as estruturas químicas de outros flavonoides glicosilados que foram desreplicados:

- Quercetina-3-*O*-metil-7-glucuronideo (269):



- Tamarixetina-3-*O*-β-(6"-malonil-glucopiranosideo):



Tendo em vista a desreplicação que foi realizada, podemos calcular a acurácia do TOF, Orbitrap e Ion-Trap, desde que a substância tenha sido encontrada nestes três analisadores, tal como representada no Quadro 5.

Fica evidente a superioridade do Orbitrap apresentando acurácia menor que 1,0 ppm (0,7 ppm em média, neste experimento), o que é de grandiosa vantagem em estudos metabolômicos com desreplicação, uma vez que a precisão em acertar a *m*/*z* é muito importante, por meio da qual podemos gerar uma lista de possíveis fórmulas moleculares e, consequentemente, pesquisarmos por fórmulas estruturais em bancos de estruturas químicas, como o *Scifinder*. O Ion-trap embora traga grandes vantagens com a fragmentação direcionada, apresentou uma acurácia inferior, da ordem de 4000 ppm em média, o que traz grande imprecisão na determinação da relação massa carga. O que foi verificado aqui em termos de acurácia está de acordo com a literatura (Holcapek *et al*., 2012). Entretanto, a acurácia e resolução dos analisadores não foi determinante neste experimento, uma vez que os três foram utilizados de maneira integrada (em conjunto).

Comparação dos analisadores dos espectrômetros de massas							
m/z exp. TOF	m/z exp. Ion-Trap	m/z exp. Orbitrap	possível substância	massa exata	Acurácia TOF	Acurácia Ion-Trap	Acurácia OBT
355.1019	355.1232	355.1027	1-caffeoylquinic acid	355.1029	3E-06	6E-05	6E-07
355.0979	354.4096	355.1023	4-caffeoylquinic acid	355.1029	1E-05	2E-03	2E-06
465.0890	465.2964	465.1031	isoquercetrin	465.1033	3E-05	4E-04	5E-07
479.0851	479.1423	479.0822	quercetrin-3-glucuronide (271)	479.0825	5E-06	1E-04	7E-07
-	499.4069	517.1350	1,4-dicaffeoylquinic acid	517.1346		3E-02	8E-07
463.0902	**463.1685**	**463.0868**	**kaempferol-3-O-glucuronide**	**463.0876**	**6E-06**	**2E-04**	**2E-06**
479.1209	479.0199	479.1189	isorhamnetin-3-O-glucoside	479.1189	4E-06	2E-04	8E-08
493.1007	**493.0981**	**493.0981**	**quercetin-3-O-methyl-7-glucuronide (269)**	**493.0982**	**5E-06**	**3E-07**	**3E-07**
517.1385	517.4185	517.1347	3,4-dicaffeoylquinic acid	517.1346	8E-06	5E-04	2E-07
535.1113	**535.3401**	**535.1090**	**kaempferol-3-O-(6''-malonyl-glucoside)**	**535.1088**	**5E-06**	**4E-04**	**3E-07**
565.1209	565.4011	565.1199	tamarixetin-3-O-β-(6''-malonyl-glucopyranoside)	565.1193	3E-06	5E-04	1E-06
				Média	8E-06	4E-03	7E-07

Quadro 5 - Informação sobre os valores de acurácia alcançados pelos analisadores TOF, Ion-Trap e Orbitrap

Assim sendo, pode-se também notar que os picos apontados pela quimiometria evidentemente aparecem no gráfico de *loadings*. Por exemplo, a substância quercetina-3-*O*-metil-7-glucuronideo (*m*/*z* 491) foi apontada tanto no modelo “MZmine TOF 1 pos” quanto no modelo “MZmine Orbitrap 3 neg” (Figura 56), neste último modelo também foi apontada a substância kaempferol-3-*O*-glucuronideo (*m*/*z* 461), todas destacadas em amarelo.

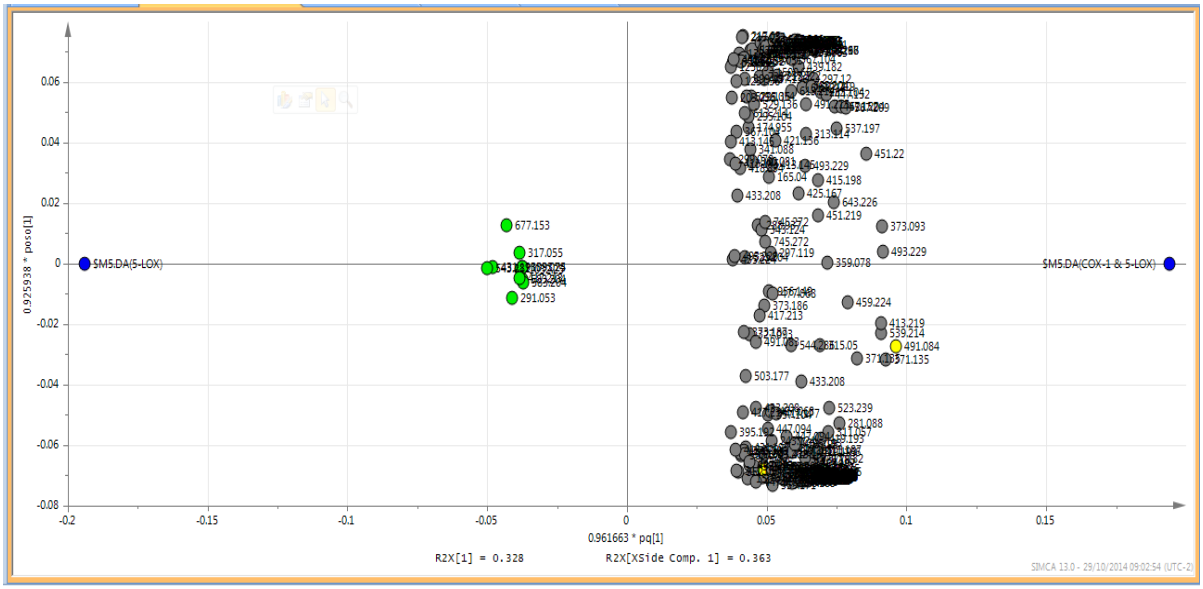


Figura 56 – Modelo “MZmine Orbitrap 3 neg” evidenciando as substâncias quercetina-3-*O*-metil-7-glucuronideo (*m*/*z* 491) e kaempferol-3-*O*-glucuronideo (*m*/*z* 461).

No gráfico de *loadings* do modelo “MZmine TOF 1 pos” também podemos notar as substâncias desreplicadas kaempferol-3-*O*-glucuronideo (*m*/*z* 463), kaempferol-3-*O*-(6"-malonil-glucosideo) (*m*/*z* 535), e tamarixetin-3-*O*-β-(6"-malonil-glucopiranosideo) (*m*/*z* 565), todos destacados em amarelo (Figuras 57 e 58). Tal como informado, todas estas substâncias são bons discriminantes da dupla inibição das enzimas COX-1 e 5-LOX.

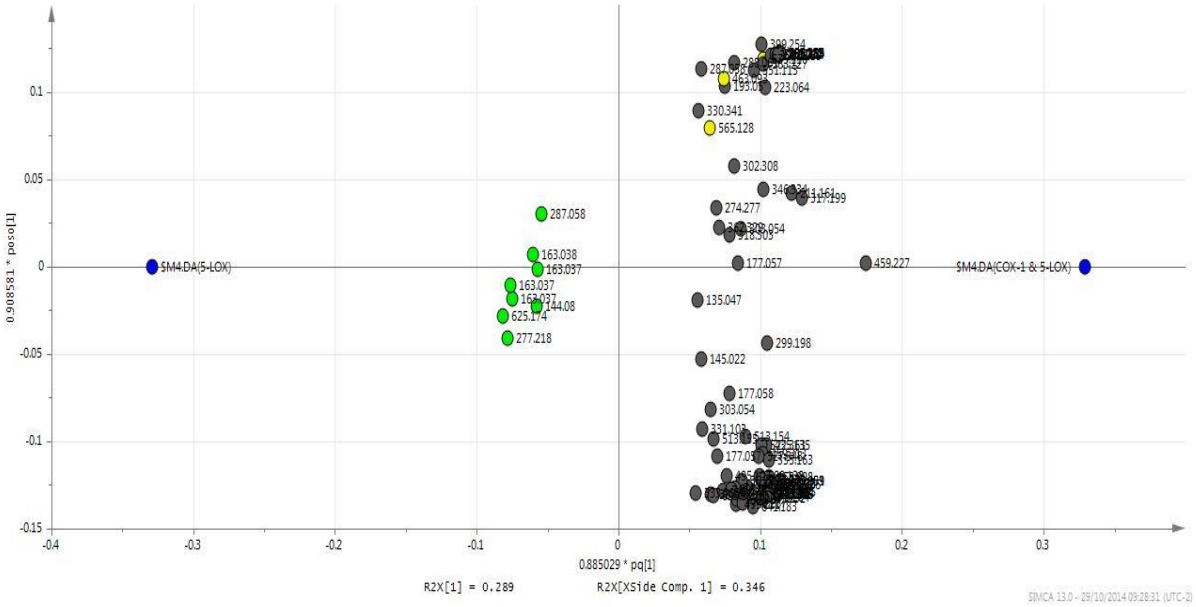


Figura 57 - Gráfico de *loadings* do modelo "MZmine TOF 1 pos" evidenciando a *m/z* substâncias de *A. trichophylla* discriminantes para a inibição de COX-1 e 5-LOX.

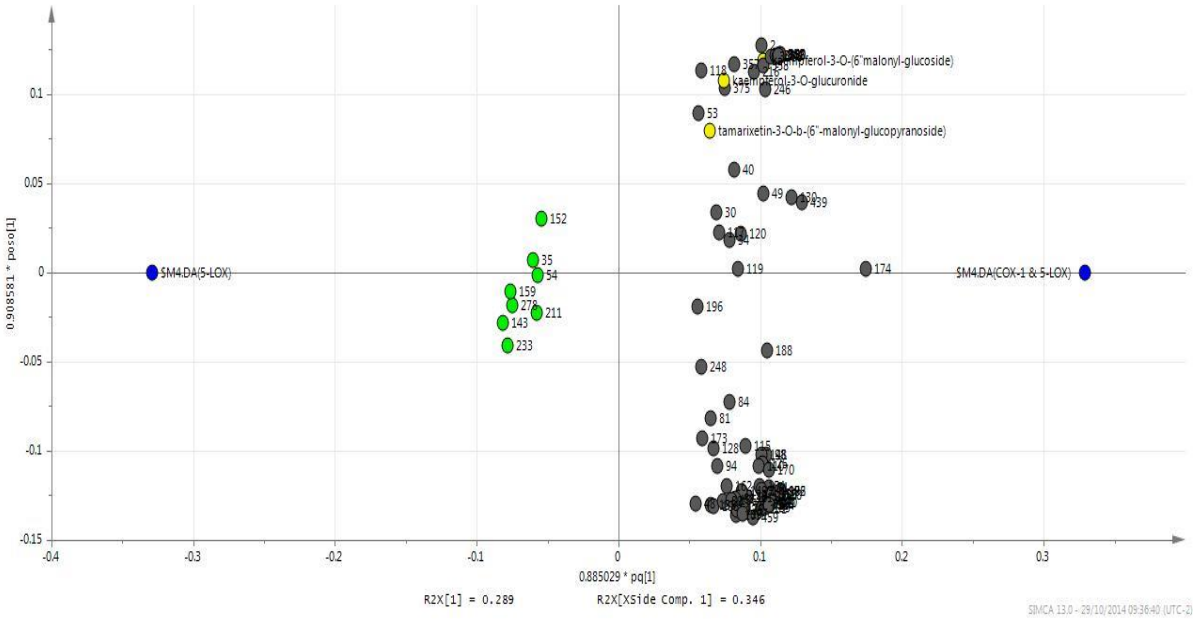


Figura 58 - Gráfico de *loadings* do modelo "MZmine TOF 1 pos" evidenciando o nome (em inglês) das substâncias de *A. trichophylla* discriminantes para a inibição de COX-1 e 5-LOX e tamarixetin-3-O-β-(6”-malonil-glucopiranosideo)

Por fim, cabe ressaltarmos aqui que flavonoides glicosilados já foram descritos na literatura (Kim *et at*., 2004; Cazarolli *et al*., 2008; Tuñón *et al*., 2009) como substâncias químicas com atividade anti-inflamatória tanto em ensaios *in vitro* como em testes *in vivo*. Além disso, outros artigos mencionam que tais substâncias atuam como inibidoras de enzimas COX e LOX (Charlier e Michaux, 2003). Entretanto, devemos notar aqui que tais substâncias

são, portanto, possivelmente ativas contra COX-1 e 5-LOX, mas não sabemos se elas possuiriam atividade contra as outras isoformas enzimáticas, uma vez que sequer tais ensaios foram realizados com os extratos brutos. Além disso, poderíamos comparar (Figura 59) tais estruturas químicas com aquelas apresentadas pela literatura (Charlier e Michaux, 2003) como inibidoras *dual*, as quais apresentam uma porção tricíclica, inibidora de COX, e uma outra porção que interfere nas vias 5-LOX dependente.

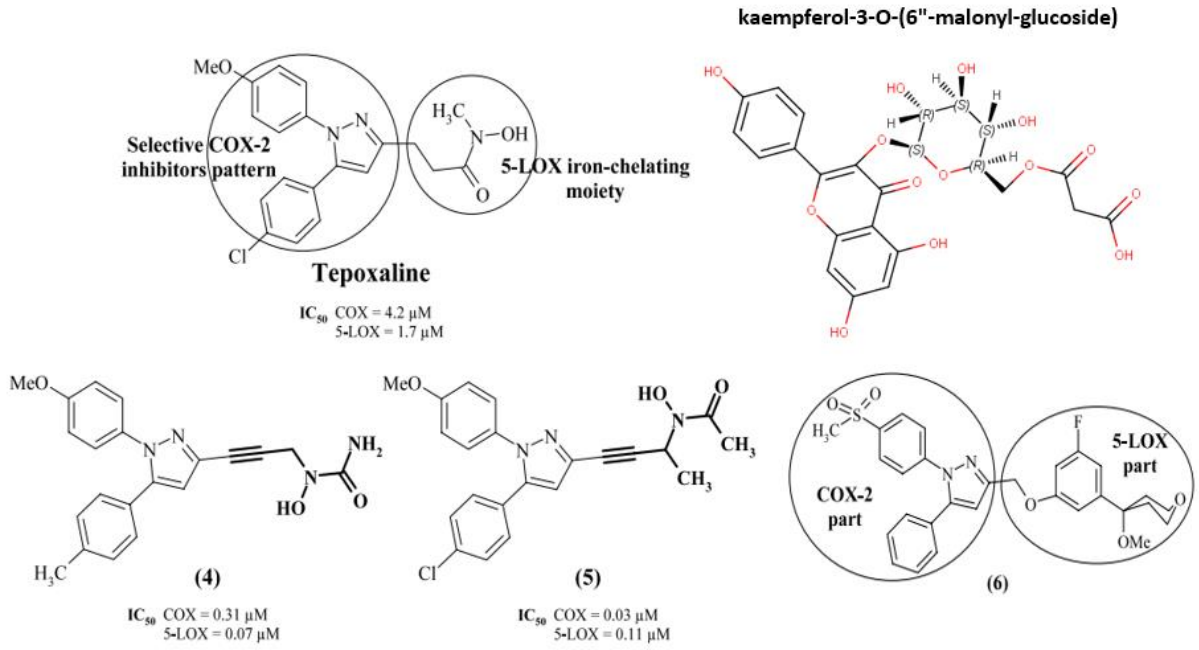


Figura 59 - Comparação estrutural de inibidores *dual* (presentes em Charlier e Michaux, 2003) e do flavonóide glicosilado.

5 – Fitoquímica direcionada das prováveis substâncias capazes de inibir COX-1 e 5-LOX.

A fitoquímica iniciou-se com o particionamento do ET, rendendo aproximadamente 1,5 g da fração diclorometano, 1,5 g da fração acetato de etila, e 13,0 g da fração aquosa. Após análises previas no HPLC-PDA-TOF e UHPLC-PDA-Orbitrap, decidiu-se trabalhar com apenas com a fração AcOEt, uma vez que seu perfil químico mostrou-se mais promissor para se realizar a desreplicação. Assim sendo, após os picos terem sido desreplicados, esta fração seguiu para o seu fracionamento em coluna com *Sephadex* LH-20, da qual se obteve 70 frações, as quais foram juntadas de acordo com a semelhança de perfil químico apresentado pela CCD analítica (Figura 60).

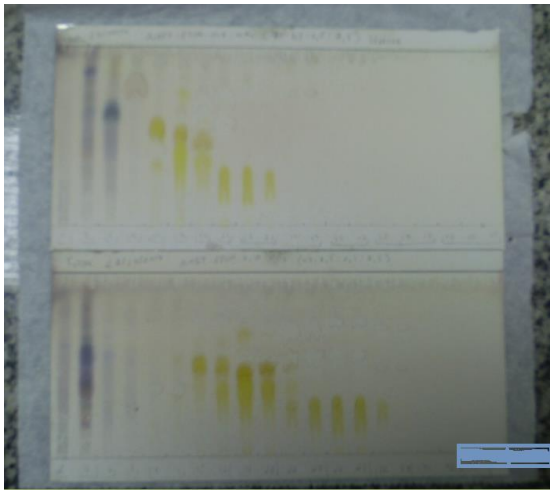


Figura 60 - CCD analítica das frações de *sephadex*

Após serem juntadas estas frações renderam 21,1 g (S-1), 21.1 g (S-3), 16.3 g (S-4), 16.2 g (S-5), 21.4 g (S-7), 20.3 g (S-12), 29.2 g (S-16), 29.5 g (S-19), 30.3 g (S-27), 29.1 g (S-37) e 30.6 g (S-45). Estas frações foram analisadas via HPLC-PDA-TOF e UHPLC-PDA-Orbitrap. Neste último a análise foi de difícil entendimento porque houve sobreposição dos picos e, portanto, optou-se por trabalhar apenas com aquela via TOF, por onde se pode escolher os picos possíveis de serem isolados (aqueles apontados pela quimiometria e/ou em concentração adequada e distantes em aproximadamente 1 minuto dos demais) via HPLC-semi-preparativo utilizando-se uma coluna C-18 monolítica de 100 x 10 mm (130 Å de poro). A análise via TOF foi realizada por meio do registro dos picos em 254 e 325 nm e do TIC ou cromatograma do íon extraído (EIC). Por exemplo, a Figura 61 mostra o registro de algumas frações em 254 e 325 nm e a Figura 62 mostra que a *m*/*z* 463 fora encontrada em 22,5 min no cromatograma do íon extraído para a fração S-16.

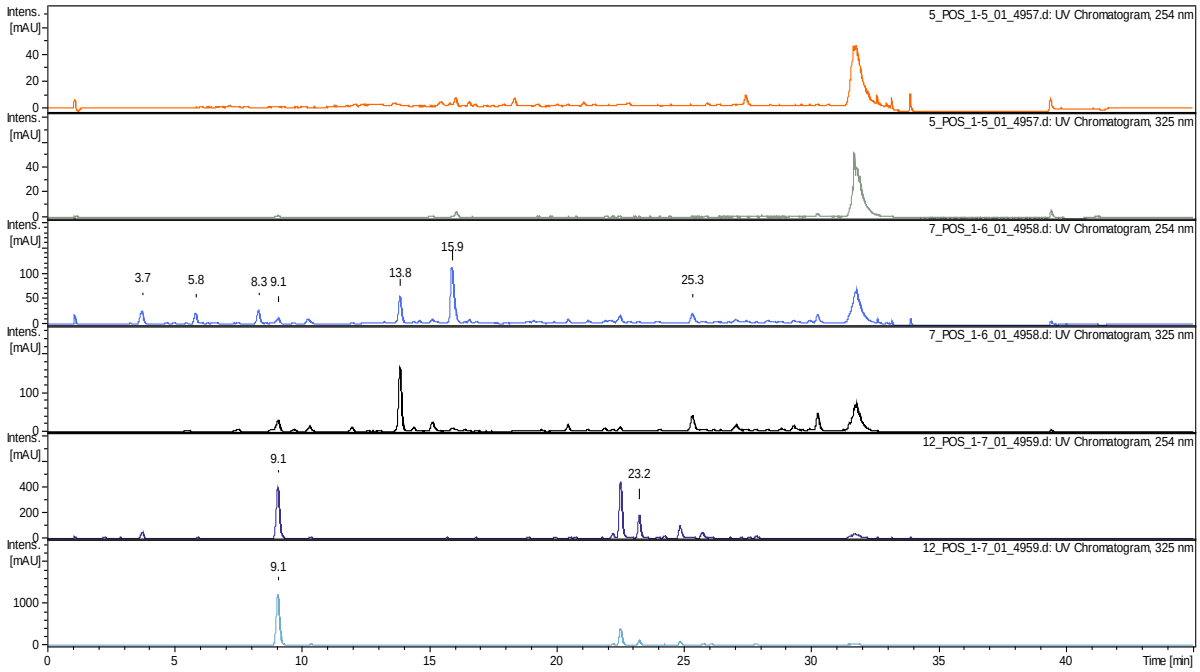


Figura 61 - Registro de algumas frações em 254 e 325 nm.

É importante mencionar que sempre é possível haver *delay* do espectro de massas em relação ao espectro no UV e, além disso, pequenas alterações no tempo de retenção foram observadas, visto que as substâncias são muito sensíveis a qualquer alteração nas concentrações da fase móvel. Ainda assim, para contornarmos estes problemas bastou-nos sermos mais flexíveis em relação ao tempo de retenção.

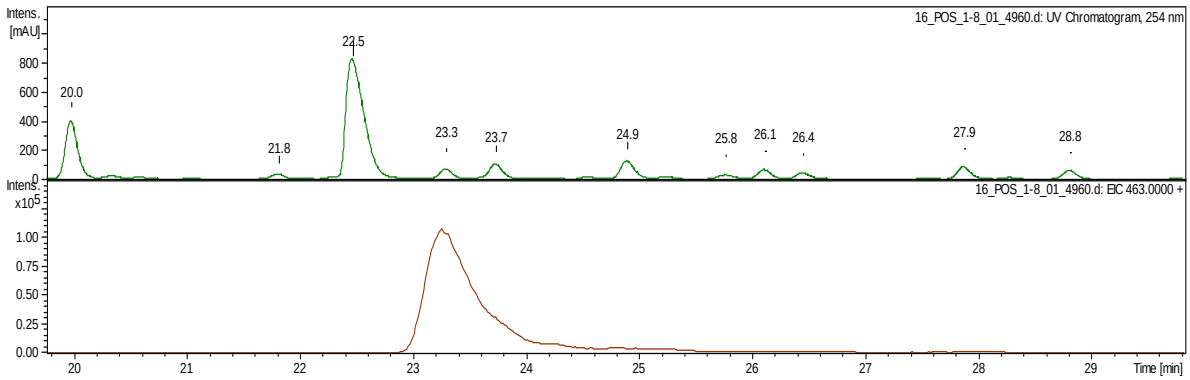


Figura 62 - EIC *m/z* 463 fora encontrada em 22,5 min no cromatograma do íon extraído para a fração S-16.

De fato, a substância em 22,5 minutos da Figura 62 corresponde ao pico de *m/z* 463, o qual fora desreplicado como kaempferol-3-*O*-glucuronideo de acordo com o perfil no UV

(Figura 63) e espectro no modo positivo (Figura 64) observados. Assim sendo, este é um exemplo de que este pico foi um dos escolhidos para isolamento.

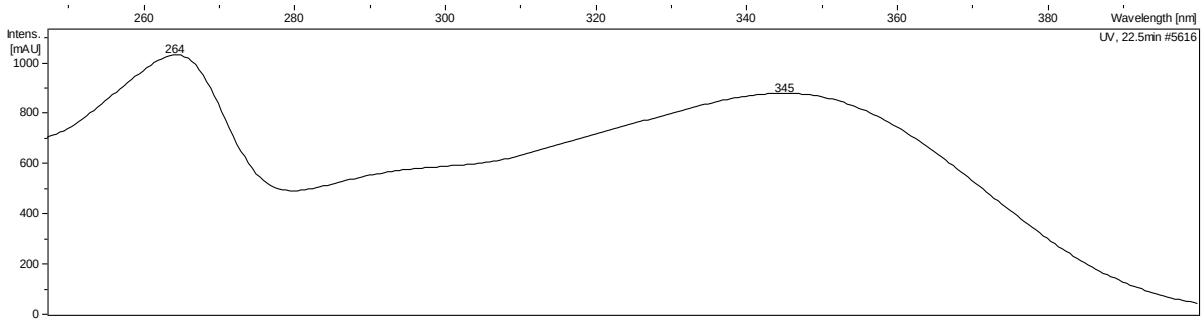


Figura 63 - Perfil de UV evidenciando se tratar da substância kaempferol-3-*O*-glucuronideo.

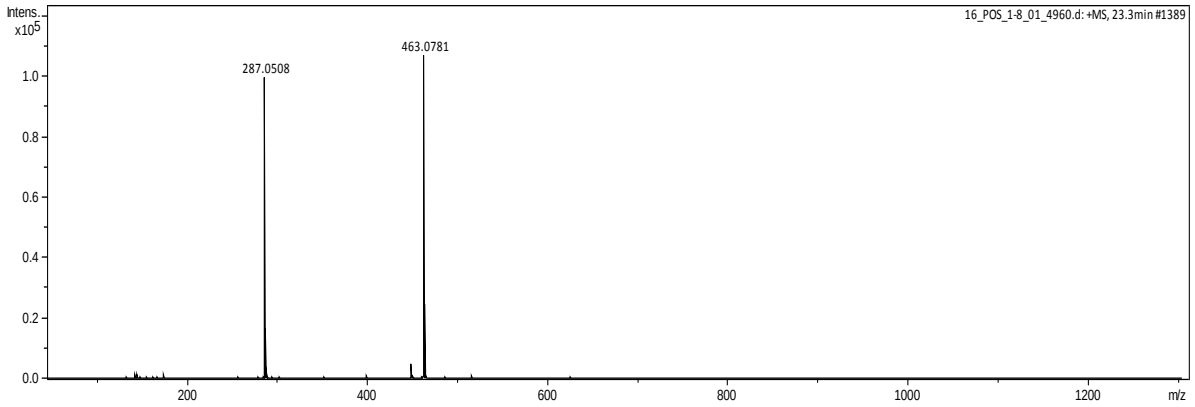


Figura 64 – Espectro de massas para a substância kaempferol-3-*O*-glucuronideo

Após as análises de todas as frações, foi possível escolher algumas das quais se poderia trabalhar com o isolamento via HPLC-semi-preparativo. Assim sendo, a estratégia adotada foi:

- isolar os picos nos tempos 9,1 min (*m*/*z* 355), 24.9 (*m*/*z* 535) e 25.7 (*m*/*z* 565) da S-12;
- isolar o pico no tempo 22,5 min (*m*/*z* 463) da S-16;
- isolar os picos 20,0 (*m*/*z* 465), 23,8 (*m*/*z* 493) e 24,5 (*m*/*z* 517), este último mais fácil na fração 27 da S-19.

Por fim, após o isolamento por HPLC-semi-preparativo dos picos escolhidos o rendimento de cada massa foi calculado e o código de cada pico isolado é o seu respectivo valor de *m*/*z* da molécula protonada de cada substância: 5 mg (código 355), 9 mg (463), 2 mg (565), 0.9 mg (465-a), 10 mg (465-b) e 7 mg (517). Após este isolamento estas substâncias

foram analisadas via RMN ^{1}H visando-se corroborar a estrutura. Por exemplo, o perfil químico observado no HPLC-semi-preparativo proveniente de uma coluna ideal para análises semi-preparativas nem sempre reproduz o observado em uma coluna monolítica, sendo que este foi o caso da substância de *m*/*z* 465, onde diferentes frações apresentaram um pico que poderia ser a substância, mas eluindo em tempos e com intensidades de absorção de UV diferentes. Sendo assim, optou-se por coletar os dois picos, nomeando-os de "465-a" e "465-b". Para resolvermos este problema bastaríamos injetar as duas substâncias em um espectrômetro de massas e confirmarmos qual delas é de fato a de *m*/*z* 465, desreplicadas como isoquercetrina. Entretanto isso não foi necessário, como explicado no próximo item.

6 – Ressonância Magnética Nuclear de Hidrogênio (RMN ^{1}H)

Para confirmar os dados experimentais de RMN ^{1}H obtidos para os picos isolados do item anterior, foi realizada a predição de dados de deslocamento químico (δ_H) e constantes de acoplamento (*J*) calculada com base no desenho das estruturas desreplicadas no *software MarvinSketch* (ChemAxon). Além dos valores preditos de RMN ^{1}H para as estruturas desreplicadas, também foram utilizados dados da literatura para confrontar com os dados experimentais. Cabe mencionar que aqui não serão demonstrados os valores de RMN ^{1}H experimentais que foram muito distintos daqueles preditos pelo *MarvinSketch*, os quais são provenientes dos picos 463, 465-a e 493. Tais picos isolados deverão ser analisados via espectrometria de massas, visando-se confirmar se, de fato, a substância foi isolada, ou se fora isolada equivocadamente uma outra substância não relevante a este estudo. Com exceção destes picos, todos os outros espectros de RMN ^{1}H apresentaram-se muito semelhantes àqueles gerados pelo *MarvinSketch*, sugerindo que, de fato, as substâncias foram desreplicadas corretamente e isoladas. Entretanto, serão demonstrados aqui apenas os espectros provenientes dos picos 355 e 535. O espectro 565 apresentou-se de intensidade muito baixa, devido à sua baixa concentração e, portanto, foi de difícil análise. Também cabe mencionar que os espectros aqui apresentados não mostram os hidrogênios das hidroxilas, uma vez que o experimento foi realizado com MeOD. Além disso, como foi isolada pouca massa, os picos de H_2O (4,87 ppm), MeOH (3,34 ppm) e HAc (1,99 ppm) estiveram presentes nas análises e os seus deslocamentos químicos podem ser observados de acordo com a literatura (Gottlieb *et al*., 1997). Além disso, o espectro do pico 465-b revelou-se mais semelhante ao espectro predito gerado pelo *MarvinSketch* para a estrutura 465 (*m*/*z* 465, isoquercetrina) do que aquele gerado pelo pico 465-a.

Deste modo, o pico isolado como 355, desreplicado como ácido 1-*O*-cafeoilquínico apresentou o seguinte espectro de RMN ^{1}H: δ 7,56 (d, *J* = 15,9 Hz, 1H, H-20), 7,05 (s, 1H, H-22), 6,96 (d, *J* = 8,2 Hz, 1H, H-13), 6,78 (d, *J* = 8,2 Hz, 1H, H-25), 6,26 (d, *J* = 15,9 Hz, 1H, H-18), 4,17 (s, 1H, H-29), 3,72 (d, *J* = 3,0 Hz, 1H, H-13), o qual pode visualizado na Figura 65. Os hidrogênios ligados aos carbonos de numeração 12 e 15 se acoplam muito e formam um multipleto no espectro experimental.

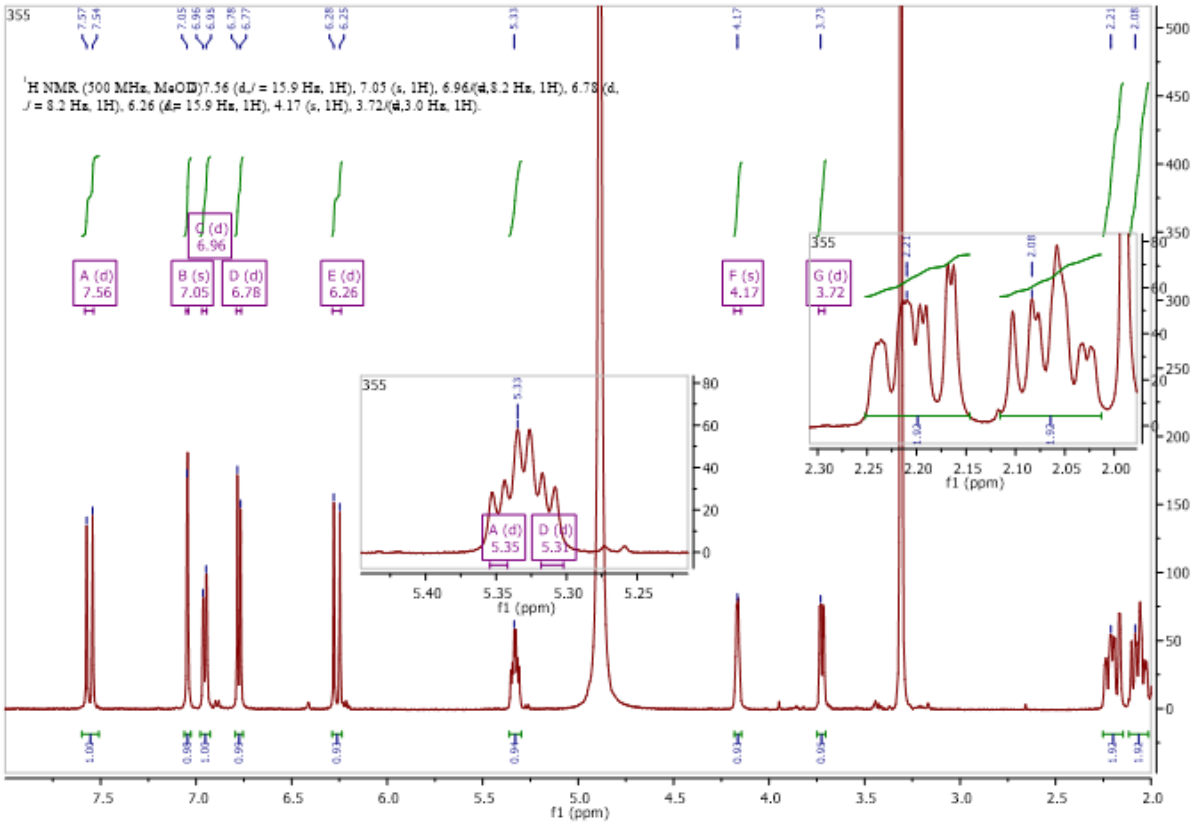


Figura 65 - Espectro de RMN ^{1}H apresentado para o pico 355.

Este espectro apresentou-se muito semelhante àquele gerado *software MarvinSketch* (Figura 66), exceto pelo primeiro singleto em δ 7,05 ppm (H-22), o qual se inverte com o dubleto em 6,96 ppm (H-13). Assim sendo, no espectro do *MarvinSketch* o deslocamento apresentado pelo hidrogênio de número 13 (H-13) apresenta um deslocamento químico maior, em 7,13 ppm. Isso sugere que, possivelmente, se trata de um isômero *cis*, o qual pode surgir de uma isomerização do isômero *trans* (Xie *et al*., 2011). Além disso, deve-se levar em consideração que os valores gerados pelo *software* são preditos, ou seja, teóricos, e não necessariamente devem ser totalmente idênticos aos do espectro experimental. Por esta

mesma razão, observamos que os valores do espectro de RMN do *Marvin* contém deslocamentos químicos mais altos em relação aos valores experimentais.

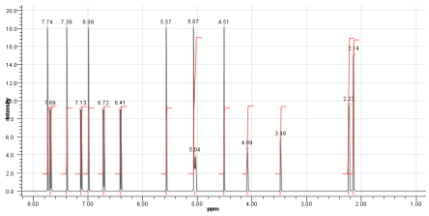


Figura 66 - Valores de deslocamento preditos para os hidrogênios do ácido 1-O-*E*-cafeoilquínico.

Os valores de deslocamento da Figura 66 para cada átomo de hidrogênio do ácido 1-O-*E*-cafeoilquínicos podem ser visualizados na Figura 67.

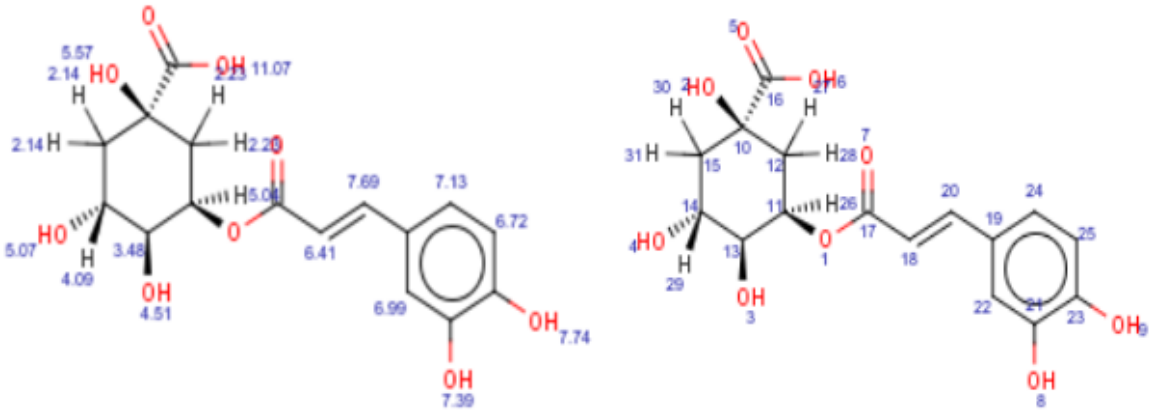


Figura 67 - Numeração dos átomos de hidrogênio e seus respectivos valores de deslocamento químico dos ácidos 1-O-*E*-cafeoilquínicos.

Já para o pico 535, desreplicado como kaempferol-3-O-(6''-malonil-glucosideo), o espectro de RMN ^{1}H experimental forneceu: δ 8,03 (d, *J* = 8,6 Hz, 2H, H-13,14), 6.88 (d, *J* = 8,7 Hz, 2H, H-12,15), 6,42 (s, 1H, H-1), 6,22 (s, 1H, H-3), 5,15 (d, *J* = 7,0 Hz, 1H, H-22),

4,23 (d, *J* = 11,7 Hz, 2H, H-29,29) (Figura 68). Este espectro, com exceção dos deslocamentos químicos dos hidrogênios das hidroxilas, os quais não aparecem no RMN ^{1}H da substância isolada 535, parece se correlacionar bastante com o espectro predito pelo *MarvinSketch* (Figura 69). Além dos deslocamentos, os acoplamentos também se correlacionam.

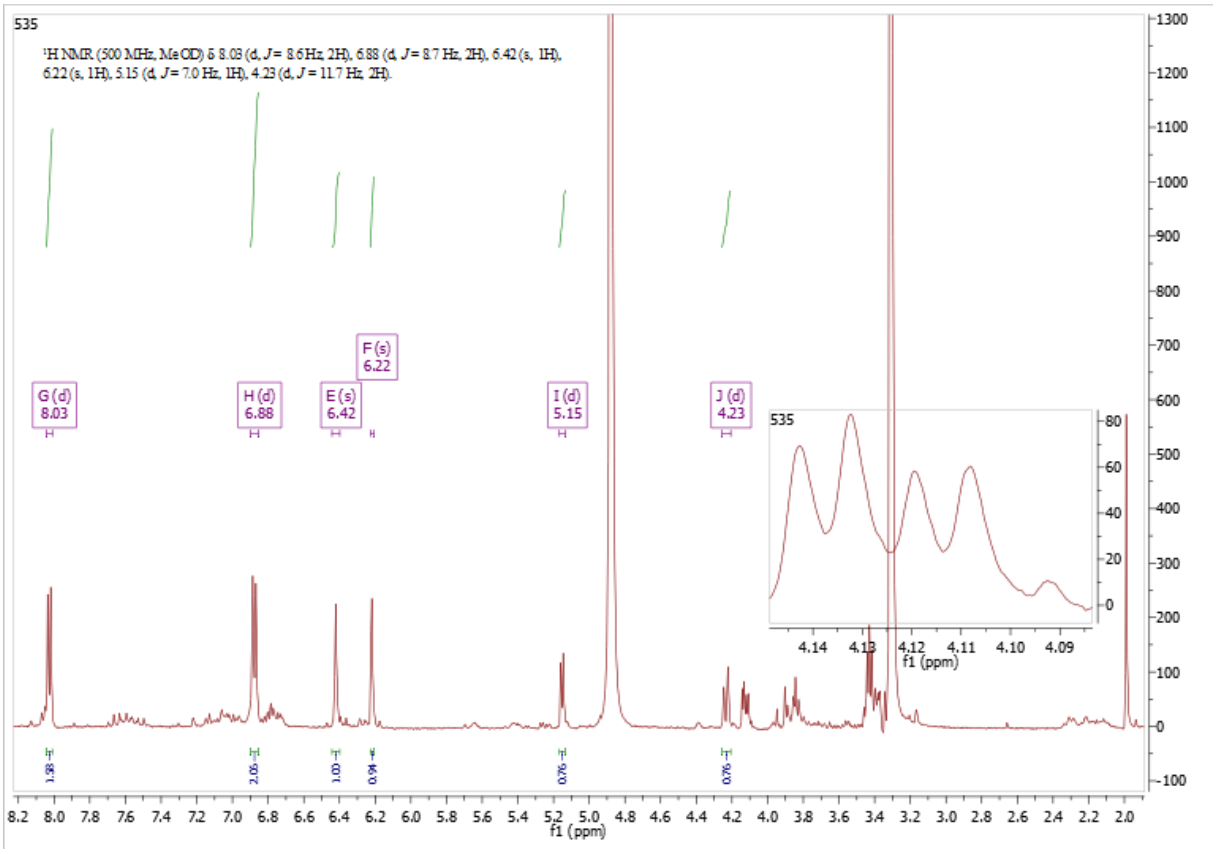


Figura 68 - Espectro de RMN ¹H apresentado para o pico 535.

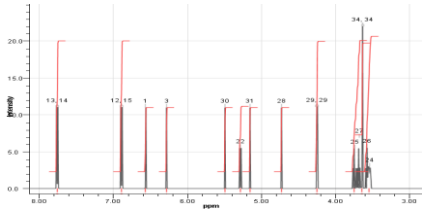


Figura 69 - Numeração dos átomos de hidrogênio e os respectivos valores de deslocamento químico do kaempferol-3-*O*-(6"-malonil-glucosideo).

Os valores de deslocamento da Figura 69 para cada átomo de hidrogênio do kaempferol-3-*O*-(6"-malonil-glucosideo) podem ser visualizados na Figura 70.

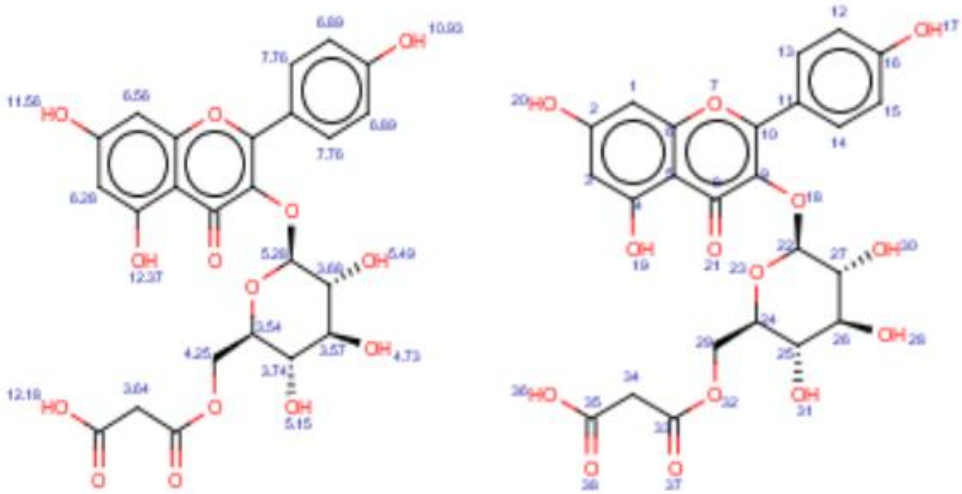


Figura 70 - Numeração dos átomos de hidrogênio e seus respectivos valores de deslocamento químico para do kaempferol-3-O-(6"-malonil-glucosideo)

Por fim, futuramente, estes dados de RMN ^{1}H serão investigados em detalhe e ainda serão obtidos os espectros bidimensionais HMBC e HSQC das substâncias isoladas, inclusive

daquelas cujos espectros não se assemelharam ao RMN ^{1}H predito pelo *MarvinSketch*, visando-se resolver, de fato, a elucidação estrutural. Por fim, ao se confirmar a pureza e elucidação estrutural, as substâncias kaempferol-3-*O*-glucuronideo, quercetina-3-*O*-metil-7-glucuronideo e kaempferol-3-*O*-(6''-malonil-glucosideo), que foram apontadas pela quimiometria como possíveis inibidoras de COX-1 e 5-LOX, deverão ter a sua atividade avaliada por meio dos ensaios *in vitro* tal como descrito em “Material e Métodos”.

Conclusão

Neste trabalho foram abordados dois assuntos distintos: quimiotaxonomia e inibição das enzimas pró-inflamatórias COX-1 e 5-LOX, os quais foram analisados com dados químicos de *Aldama* La Llave provenientes da abordagem metabolômica não-direcionada e tratados por meio de análises quimiométricas. Deste modo, esta conclusão pode-se se subdividir, do ponto de vista lógico, em dois grandes tópicos, cada qual tratando-se de um assunto diferente.

Primeiramente, quanto às contribuições quimiotaxonômicas, tendo em vista a metodologia empregada, foi possível verificar que o gênero *Aldama* La Llave é quimicamente consistente, formando *clusters* bem definidos em relação aos demais gêneros analisados. Tal fato foi mais observado por meio da HCA que se utilizou de dados provenientes dos modos de ionização positivo e negativo concomitantemente. A HCA pode explicar alguns agrupamentos infragenéricos para determinadas espécies que apresentam problemas de classificação taxonômica com base em dados de morfoanatomia. Futuramente, esta HCA poderá ser comparada com os dados de filogenia baseada em dados moleculares. Além disso, pouca correlação pôde-se observar entre os dados químicos e a classificação ambiental de diversas regiões do Brasil, bem como a classificação de Blake (1918), por meio de todas as PCA. Entretanto, diversos modelos de predição com base em OPLS-DA foram apresentados nesta Dissertação, revelando o grande potencial desta técnica em predizer amostras desconhecidas ou externas ao grupo de estudo. Deste modo, foi gerada uma lista de picos de íons, detectados pelo UHPLC-Orbitrap, como sendo discriminantes de *Aldama* La Llave. Tais picos foram correlacionados com um banco de massas criado para substâncias químicas específicas do gênero *Aldama* e outra proveniente de Asteraceae, por meio dos quais foi possível encontrar 37 prováveis substâncias químicas. Contudo, tais substâncias químicas não puderam ser desreplicadas com base no perfil de UV e padrão de fragmentação, visto que muitas das quais ou estavam em baixa intensidade no espectro, ou possuíam um confuso padrão de fragmentação visualizado pelo espectro de MS-MS. Foi comentado que tal dificuldade da análise de fragmentação possa se dever ao fato do Orbitrap Exactive Plus não conter um analisador Quadrupolo, o qual poderia selecionar íons específicos para a desreplicação, ou ainda o fato do UHPLC não estar conseguindo separar bem os picos de acordo com a análise empregada. Ainda assim, em um futuro próximo, tais substâncias terão seus tempos de retenção informados por meio de modelos experimentais de predição o que, certamente, favorecerá uma desreplicação mais inequívoca destas substâncias.

Quanto às análises realizadas para predição de prováveis substâncias químicas responsáveis pela inibição dupla das enzimas pro-inflamatórias COX-1 e 5-LOX, primeiramente foi listado que apenas *A. trichophylla* e *A. robusta* apresentaram a inibição dupla desejada. A primeira espécie foi escolhida para a fitoquímica direcionada, uma vez que ela é menos estudada do ponto de vista fitoquímico. Para o intuito de se saber quais seriam os prováveis picos responsáveis pela inibição dupla e, presentes em *A. trichophylla*, uma série de modelos experimentais com base nos diferentes analisadores TOF e Orbitrap foram criados empregando-se diferentes combinações de algoritmos do *software MZmine*, bem como os *softwares MetAlign* e *MSClust*. Tais comparações são inéditas na literatura para o propósito aqui abordado. Assim sendo, dentro do material e das técnicas utilizadas, foi constatado que os algoritmos do *MZmine* tinham pouca influência na melhora do modelo experimental, assim como os *softwares MetAlign* e *MSClust*. Entretanto, de uma maneira geral, os dados químicos provenientes do modo de ionização negativo pareceram um pouco melhores na criação dos modelos de inibição enzimática. Ainda assim, foram escolhidos seis melhores modelos (três provenientes de dados do TOF e três do Orbitrap) para serem otimizados por meio da deleção de variáveis não importantes na projeção (VIP < 1). Tais modelos também foram rodados no *software* Weka, por meio do qual foi realizada a *Decision Tree*, informando quais seriam os picos mais importantes na classificação do modelo. Com base nisso, os modelos otimizados e a *Decision Tree* forneceram uma série de prováveis substâncias responsáveis pela inibição dupla das enzimas COX-1 e 5-LOX, resultando em apenas três substâncias desreplicadas com base no perfil de UV e padrões de fragmentação visualizados por meio de análises realizadas no HCD-Orbitrap e Ion-Trap: kaempferol-3-*O*-glucuronideo, quercetina-3-*O*-metil-7-glucuronideo e kaempferol-3-*O*-(6"-malonil-glucosideo). Estas três substâncias foram isoladas de frações em coluna de *Sephadex LH-20* da fração AcOEt de *A. trichophylla* e uma prévia análise pelo RMN ^{1}H foi realizada, indicando que, possivelmente, foi obtido sucesso no isolamento. Por fim, a análise de elucidação estrutural será melhor concluída e as substâncias poderão ter o seu potencial de inibição testado por meio das análises *in vitro* com COX-1 e 5-LOX.

Referências

Alvarenga, S.A.V.; Ferreira, M.J.P.; Emerenciano, V.P.; Cabrol-Bass, D. Chemosystematic studies of natural compounds isolated from Asteraceae: characterization of tribes by principal component analysis. Chemometric and Intelligent Laboratory Systems, v. 56, p. 27-37, 2001.

Alvares, C.A.; Stape, J.L.; Sentelhas, P.C.; de Moraes, G.; Leonardo, J.; Sparovek, G. Köppen's climate classification map for Brazil. Meteorologische Zeitschrift, v. 22, p. 711-728, 2013.

Ambrosio, S.R.; Schorr, K.; Da Costa, F.B. Terpenoids of *Viguiera* arenaria (Asteraceae). Biochemical Systematics and Ecology, v. 32, p. 221–224, 2004.

Ambrosio, S.R.; Tirapelli, C.R.; Bonaventura, D.; Oliveira, A.M.; Da Costa, F.B. Pimarane diterpene from *Viguiera* arenaria (Asteraceae) inhibit rat carotid contraction. Fitoterapia, v. 73, p. 484–489, 2002.

Bhargava, N.; Sharma, G.; Bhargava, R.; Mathuria, M. Decision tree analysis on J48 algorithm for data mining, v. 3, p. 1114-1119, 2013.

BLAKE, S.F. A revision of the genus *Viguiera*. Contributions of the Gray Herbarium, v. 54, p. 1-205, 1918.

Blaker, J.G. Compositae IV: Helianthoideae. In C.F.P Martius & A.W. Eichler (eds) Flora brasiliensis Monacchii. Typografia Regia, v. 6, p. 135-268, 1884.

Buydens, L.M.C.; Reijmers, T.H.; Beckers, M.L.M.; Wehrens, R. Molecular data-mining: a chellenger for chemometrics. Chemometrics and Intelligent Laboratory Systems, v. 49, p. 121-133, 1999.

Canales, M. *et al*. Antimicrobial Activity of the Extracts and Essential Oil of *Viguiera dentata*. Pharmaceutical Biology, v. 46, p. 719–723, 2008.

Castillo, S.; Gopalacharyulu, P.; Yetukuri, L.; Orešič, M. Algorithms and tools for the preprocessing of LC–MS metabolomics data. Chemometrics and Intelligent Laboratory Systems, v. 108, p. 23–32, 2011.

Cazarolli, L.H. *et al*. Flavonoids: prospective drug candidates. Mini-Reviews in Medicinal Chemistry, v. 8, p. 1429-1440, 2008.

Chagas-Paula, D. A. *et al*. Chlorogenic acids from *Tithonia diversifolia* demonstrate better anti-inflammatory effect than indomethacin and its sesquiterpene lactones. Journal of ethnopharmacology, v. 136, p. 355–362, 2011.

Charlier, C. & Michaux, C. Dual inhibition of cyclooxygenase-2 (COX-2) and 5-lipoxygenase (5-LOX) as a new strategy to provide safer non-steroidal anti-inflammatory drugs. European Journal of Medicinal Chemistry, v. 38, p. 645–659, 2003.

ChemAxon – Cheminformatic Platforms and Desktop. Disponível em: www.chemaxon.com.

Clifford, M.N.; Johnston, S.K.; Knight, S.; Kuhnert, N. Hierarchical scheme for LC-MS^n identification of chlorogenic acids. Journal Agricultural and Food Chemistry, v. 51, p. 2900-2911.

Clifford, M.N.; Knight, S.; Kuhnert, N. Discriminating between the six isomers of dicaffeoylquinic acid by LC-MS^n. Journal of Agricultural and Food Chemistry, v. 53, p. 3821-3832.

Da Costa, F. B.; Albuquerque, S.; Vichnewski, W. Diterpenes and synthetic derivatives from *Viguiera aspillioides* with trypanomicidal activity. Planta medica, v. 62, p. 557–559, 1996.

Da Costa, F.B. Asteraceae: Química e Potencial Farmacológico. 244. Texto sistematizado para livre docência. Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2008.

Dictionary of Natural Products. Disponível em: <http://dnp.chemnetbase.com/intro/>.

Domon, B. & Costello, C. A systematic nomenclature for carbohydrate fragmentation in FAB-MS/MS spectra of glycoconjugates. Glycoconjugate Journal, v. 5, p. 397-409, 1988.

Emerenciano, V.P.; Cabrol-Bass, D.; Ferreira, M.J.P.; Alvarenga, S.A.V.; Brant, A.J.C.; Scotti, M.T.; Barbosa, K.O. Chemical evolution in the Asteraceae. The oxidation-reduction mechanism and production of secondary metabolites. Natural Product Communications, v. 1, p. 495-507, 2006.

Equipe Científica AsterbioBioChem. Disponível em: <www.asterbiochem.org>.

Eugster, P.J.; Boccard, J.; Debrus, B.; Bréant, L.; Wolfender, J-L.; Martel, S.; Carrupt, P-A. Retention time prediction for dereplication of natural products ($C_xH_yO_z$) in LC-MS metabolite profiling. Phytochemistry, v. 108, p. 196-207, 2014.

Fiehn, O. Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. Comparative and functional genomics, v. 2, p. 155–168, 2001.

Fiehn, O. Metabolomics-the link between genotypes and phenotypes. Plant molecular biology, v. 48, p. 155–71, 2002.

Fiorucci, S.; Meli, R.; Bucci, M.; Cirino, G. Dual inhibitors of cyclooxygenase and 5-lipoxygenase. A new avenue in anti-inflammatory therapy? Biochemical pharmacology, v. 62, p. 1433–1438, 2001.

Funk, V.A.; Susanna, A.; Stuessy, T.F.; Bayer, R.J. Systematics, evolution, and biogeography of Compositae. International Association for Plant Taxonomy. Vienna, 2009.

Gasteiger J. & Engel, T. (Eds.), Chemoinformatics. WILEY-VCH Verlag GmbH & Co. KgaA, Weinheim, 2003.

Gobbo-Neto, L. *et al*. Evaluation of the anti-inflammatory and antioxidant activities of di-C-glucoflavones from *Lychnophora ericoides* (Asteraceae). Planta medica, v. 71, p. 3–6, 2005.

Goodacre, R.; Vaidyanathan, S.; Dunn, W.B.; Harrigan, G.G.; Kell, D.B. Trend in Biotechnology, v. 22, p. 245-252. 2004.

Gottlieb, O.R. Biodiversidade: Uma teoria molecular. Química Nova. v, 15, p. 167-172, 1992.

Gottlieb, O.R. & Borin, M.R.M.B. Químico-Biologia quantitativa: Um novo paradgma? Química Nova, v. 35, p. 2105-2114, 2012.

Hall, M.; Frak, E.; Holmes, G.; Pfahinger, B.; Reutemann, P.; Witten, I.H. The weka data mining software: An update, v. 11, p. 10-18, 2009.

Henchoz, Y.; Guillarme, D.; Martel, S.; Rudaz, S.; Veuthey, J-L.; Carrupt, P-A. Fast log P determination by ultra-high-pressure-liquid chromatography coupled with UV and mass spectrometry detections. Analytical and Bioanalytical Chemistry, v. 394, p. 1919-1930, 2009.

Holčapek, M.; Jirásko, R.; Lísa, M. Recent developments in liquid chromatography-mass spectrometry and related techniques. Journal of chromatography. A, v. 1259, p. 3–15, 2012.

Hu, Q.; Noll, R.J.; Li, H.; Makarov, A.; Hardman, M.; Cooks, G. The Orbitrap: a new mass spectrometer, v. 40, p. 430-443, 2005.

Kachlicki, P.; Einhorm, J.; Muth, D.; Kerhoas, L.; Stobiecki, M. Evaluation of glycosylation and malonylation patterns in flavonoid glycosides during LC/MS/MS metabolite profiling. Journal of Mass Spectrometry, v. 43, p. 572-586, 2008.

Katajamaa, M.; Miettinen, J.; Oresic, M. MZmine: toolbox for processing and visualization of mass spectrometry based molecular profile data. Bioinformatics, v. 22, p. 634–636, 2006.

Kim, H.P.; Son, K.H.; Chang, H.W.; Kang, S.S. Anti-inflammatory plant flavonoids and cellular action mechanisms. Journal of Pharmacological Sciences, v. 96, p. 229-245, 2004.

Larose, D.T. Data Mining – Methods and Models. Central Connecticut State University, 2006.

Lommen, A. & Kools, H.J. MetAlign 3.0: performance enhancement by efficient use of advances in computer hardware. Metabolomics, v. 8, p. 719-726, 2012.

Lommen, A. MetAlign: Interface-Driven, versatile metabolomics tool for hyphenated full-scan mass spectrometry data preprocessing. Analytical Chemistry, v. 81, p. 3079-3086, 2009.

Magenta, M.A.G. *Viguiera* Kunth (Asteraceae , Heliantheae) na América do Sul e sistemática das espécies do Brasil. Tese de Doutorado. 339. Universidade de São Paulo, 2006.

Magenta, M.A.G & Pirani, J.R. Novidades taxonômicas em Aldama (Asteraceae-Heliantheae). Rodriguesia, v. 65, p. 175-192, 2014.

Markham, K.R. Techniques of flavonoid identification. London: Academic press, v. 31, 1982.

Mutihac, L. & Mutihac, R. Mining in chemometrics. Analytic Chimica ACTA, v. 612, p. 1-18, 2008.

Ngo, L.T.; Okogun, J.I.; & Folk, W.R. 21St Century Natural Product Research and Drug Development and Traditional Medicines. Natural product reports, v. 30, p. 584–92, 2013.

Nicolete, R. *et al*. Budlein A from *Viguiera robusta* inhibits leukocyte-endothelial cell interactions, adhesion molecule expression and inflammatory mediators release. Phytomedicine, v. 16, p. 904–915, 2009.

Oliver, S.G.; Wilson, M.K.; Kell, D.B.; Baganz, F.; Systematic functional analysis of the yeast genome. Focus, v. 16, p. 373-378, 1998.

Parente, L. Pros and Cons of Selective Inhibition of Cyclooxygenase-2 versus Dual Lipoxygenase / Cyclooxygenase Inhibition : Is Two Better than One ? The Journal of Rheumatology, v. 28, p. 2375–2382, 2001.

Pluskal, T., Castillo, S., Villar-Briones, A. & Oresic, M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC bioinformatics, v. 11, p. 395–406, 2010.

Quinlan, J. R. C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, 1993.

Quinlan, J.R. Induction of decision trees. Machine Learning, v. 1, p. 81-106, 1986.

Quinlan, J.R. Simplifying decision trees. International Journal Man-Machine Studies, v. 27, p. 221-234, 1987.

Rang, H.P.; Dale, M.M.; Ritter, J.M.; Moore, P.K. Farmacologia. Elsevier Editora Ltda, Rio de Janeiro. Tradução da 5ª Edição Americana, 2004.

Rencher, A. C., Methods of Multivariate Analysis. Bringhan Young University, 2nd Ed., 2002.

Robinson, H.; Moore, A.J. New species and new combinations in *Rhysolepis* (Heliantheae: Asteraceae). Proceedings of Bioolgical Society of Washington, v. 117, p. 423-446, 2004.

Sakakibara, H.; Honda, Y.; Nakagawa, S.; Ashida, H.; Kanazawa, K. Simultaneous determination of all polyphennols in vetables, fruits, and teas. Journal of Agricultural and Food Chemistry, v. 51, p. 571-581.

Schilling, E. E. & Panero, J. L. 2011. A revised classification of subtribe Helianthinae (Asteraceae: Heliantheae) II. Derived lineages. Botanical Journal of the Linnean Society, v. 167, p. 311–331.

Schilling, E.E. & Jansen, R.K. 2012. Restriction Fragment Analysis of Chloroplast DNA and the Systematics of *Viguiera* and Related Genera (Asteraceae: Heliantheae). American Journal of Botany, v. 76, p. 1769–1778.

Schilling, E.E. & Panero, J.L. 2002. A revised classification of subtribe Helianthinae (Asteraceae: Heliantheae). I. Basal lineages. Botanical Journal of the Linnean Society, v. 140, p. 65–76.

Schilling, E.E.; Da Costa, F.B.; Lopes, N.P.; Heise, P.J. Brazilian species of *Viguiera* (Asteraceae) exhibit low levels of ITS sequence variation. Edinburgh Journal of Botany, v. 57, p. 323–332, 2000.

Schimidt, T.J. Toxic activities of sesquiterpene lactones: structural and biological aspectes. Current Organic Chemistry, v. 3, p. 577–608, 1999.

Schorr, K.; García-Piñeres, A.J.; Siedle, B.; Merfort, I.; Da Costa, F.B. Guaianolides from *Viguiera gardneri* inhibit the transcription factor NF-kappaB. Phytochemistry, v. 60, p. 733–40, 2002.

Schorr, K.; Rott, A.; Da Costa, F.B. Merfort, I. Optimisation of a human neutrophil elastase assay and investigation of the effect of sesquiterpene lactones. Biologicals, v. 33, 175–184, 2005.

Schwab, W. Metabolome diversity: too few genes, too many metabolites? Phytochemistry, v. 62, p. 837–849, 2003.

SciFinder Scholar™, version 2014. American Chemical Society: Columbus, Ohio, EUA. Disponível em: <https://scifinder.cas.org>.

Seaman, F.C. Sesquiterpene lactones as taxonomic characters in the Asteraceae. The Botanic Review, v. 48, p. 121-529, 1982

Sharma, O.P. Plant taxonomy. Tata McGraw-Hill Publishing Company Limited, New Delhi, 7ed, p. 95, 2007.

Siedle, B. *et al*. Quantitative structure-activity relationship of sesquiterpene lactones as inhibitors of the transcription factor NF-kappaB. Journal of medicinal chemistry. v. 47, p. 6042–6054, 2004.

Sokal, R.R. Numerical Taxonomy. Scientific American. v. 215, p. 106-116, 1966.

Spring, O. & Schilling, E.E. The sesquiterpene lactone chemistry of *Helianthus* sect. Atrorubentes (Asteraceae: Heliantheae). Biochemical Systematics and Ecology, v. 19, p. 59–79, 1991.

Spring, O. & Schillingt, E.E. Chemosystematic Investigation of the Annual Species of *Helianthus* (Asteraceae)*. Biochemical Systematics and Ecology, v. 17, p. 519–528, 1989.

Spring, O. & Schillingt, E.E. The Sesquiterpene Lactone Chemistry of *Helianthus* Sect . Ciliares (Asteraceae: Heliantheae)*. Biochemical Systematics and Ecology, v. 18, p. 139–143, 1990.

Spring, O.; Zipper, R.; Reeb, S.; Vogler, B.; Da Costa, F.B. Sesquiterpene lactones and a myoinositol from glandular trichomes of *Viguiera quinqueremis* (Heliantheae; Asteraceae). Phytochemistry, v. 57, p. 267–72, 2001.

Summer, L.W.; Mendes, P.; Dixon, R.A. Plant metabolomics: large-scale phytochemistry in the functional genomics era. Phytochemistry, v. 62, p. 817-836, 2003.

Taylor, P.G.; Dupuy Loo, O.A.; Bonilla, J.A; Murillo, R. Anticancer activities of two sesquiterpene lactones, millerenolide and thieleanin isolated from *Viguiera sylvatica* and Decachaeta thieleana. Fitoterapia, v. 79, p. 428–32, 2008.

Tellinghuisen, J. & Bolster, C.H. Using R^2 to compare least-squares fit models: When it must fail. Chemometrics and Intelligent Laboratory Systems, v. 105, p. 220-222, 2011.

Tikunov, Y.M.; Laptenok, S.; Hall, R.D.; Bovy, A.; de Vos, R.C.H. MSClust: a tool for unsupervised mass spectra extraction of chromatography mass spectrometry ion-wise aligned data. Metabolomics, v. 8, 714-718, 2012.

Tirapelli, C.R.; Ambrosio, S.R.; Da Costa, F.B.; De Oliveira, A.M. Evidence for the mechanisms underlying the effects of pimaradienoic acid isolated from the roots of *Viguiera arenaria* on rat aorta. Pharmacology, v. 70, p. 31–38, 2004.

Tsimogiannis, D.; Samiotaki, M.; Panayotou, G.; Oreopoulou, V. Characterization of flavonoid subgroups and hydroxyl substitution by HPLC-MS/MS. Molecules, v. 12, p. 593-606, 2007.

Tuane, O.; Bombo, A.B.; Apezzato-Da-Gloria, B. Anatomy of vegetative organs with an emphasis on the secretory structures of two species of *Aldama* (Asteraceae-Heliantheae). Botany, v. 91, p. 335-342, 2013.

Tuñón, M.J.; García-Mediavilla, M.V.; Sánchez-Campos, S.; Gonzáles-Gallego, J. Potential of flavonoids as anti-inflammatory agentes: modulation of pro-inflammatory gene expression and signal transduction pathways. Current Drug Metabolism, v. 10, p. 256-271, 2009.

Verdi, L.G.; Brighente, I.M.C.; Pizzolatti, M.G. Gênero *Baccharis* (Asteraceae): aspectos químicos, econômicos e biológicos. Química Nova, v. 28, p. 85-94, 2005.

Verpoorte, R. Exploration of nature's chemodiversity: the role of secondary metabolites as leads in drug development. Drug Discovery Today, v. 3, p. 232–238, 1998.

Vesscchi, R.; Lopes, N.P.; Gozzo, F.C.; Dörr, F.A.; Murgu, M.; Lebre, D.T.; Abreu, R.; Bustillos, O.V.; Riveros, J.M. Nomenclaturas de espectrometria de massas em língua portuguesa. Química Nova, v. 34, p. 1875-1887, 2011.

Villas-Bôas, S.G. & Gombert, A.K. Análise do Metaboloma. Biotecnologia Ciência e Desenvolvimento, v. 36, p. 58–69, 2006.

Villas-Bôas, S.G.; Rasmussen, S.; Lane, G.A. Metabolomics or metabolite profiles? Trends in biotechnology, v. 23, p. 385–386, 2005.

Vukics, V. & Guttman, A. Structural characterization of flavonoid glycosides by multi-stage mass spectrometry. Mass Spectrometry Reviews, v. 29, p. 1-16, 2008.

Witten, I.H.; Eibe, F.; Hall, M.A. Data Mining – practical machine learning tools and techniques. 3rd. Ed., Morgan Kaufmann Publishers. 2011.

Wolfgang, P. & Steinwedel, H. Ein neues massenspektrometer ohne magnetfeld. Zeitschrift Naturforschung Teil A, v. 8, p. 448, 1953.

Xie, Cen *et al*. Investigation of isomeric transformations of chlorogenic acid in buffers and biological matrixes by ultraperformance liquid chromatography coupled with hybrid quadrupole/ion mobility/orthogonal acceleration time-of-flight mass spectrometry. Journal of agricultural and food chemistry, v. 59, p. 11078-11087, 2011

Yuliana, N. D.; Khatib, A.; Verpoorte, R.; Choi, Y.H. Comprehensive extraction method integrated with NMR metabolomics: a new bioactivity screening method for plants, adenosine A1 receptor binding compounds in Orthosiphon stamineus Benth. Analytical Chemistry, v. 83, p. 6902–6906, 2011.

Zdero, C.; Bohlmann, F. Systematics and evolution within the Compositae seen with the eyes of a chemist. Plant Systematics and Evolution, v. 171, p. 1-14, 1990.

Zhou, B.; Xiao, J.F.; Tuli, L.; Ressom, H.W. LC-MS-based metabolomics. Molecular Systems, v. 8, p. 470-481, 2012.

APÊNDICE A - Parâmetros 1 do *MZmine 2.10*.

Raw data import

Peak detection

Mass detection

Mass detector: Exact mass

Noise leve: 1.0E5

FTMS shoulder peaks filter

Mass resolution: 70,000

Peak model function: Lorentzian extended

Chromatogram builder

Min time span (min): 0.3

Min heigh: 5.0E5

m/z tolerance: 0.002 or 5.0 ppm

Chromatogram deconvolution

Algorithm: Local minimum search

Chromatographic threshold: 5%

Search minimum in RT range (min): 0.2

Minimum relative height: 15%

Minimum absolute height: 5.0E5

Min ratio of peak top/edge: 5

Peak duration range (min): 0.3-10.0

Isotopes – Isotopic peaks grouper

m/z tolerance: 0.002 or 5.0 ppm

Retention time tolerance: 0.2 absolute

Maximum charge: 2

Representative isotope: Most intense

Alignment – Join aligner

m/z tolerance: 0.002 or 5.0 ppm

Weight for *m/z*: 15

Retention time tolerance: 0.2 absolute

Weight for RT: 15

Comparative isotope pattern: Isotope *m/z* tolerance: 0.002 or 5.0 ppm

Minimum absolute height: 5.0E5

Minimum score: 65%

Gap filling – Peak finder

Intensity tolerance: 1%

m/z tolerance: 0.002 or 5 ppm

Retention time tolerance: 0.2 absolute

RT correction

Identification

Complex search

Ionization method: $[M+H]^+$ for positive and $[M-H]^-$ for negative

Retention time tolerance: 0.2 absolute

m/z tolerance: 0.002 or 2.5 ppm

Max complex peak height: 50%

Adduct search

For positive ionization mode	For negative ionization mode
[M+Na-H] 21.9825 *m/z*	[M+Cl+H] 35.9774 *m/z*
[M+K-H] 37.9559 *m/z*	$[M+H_3PO_4]$ 97.9769 *m/z*
[M+Mg-2H] 21.9694 *m/z*	
$[M+NH_3]$ 17.0265 *m/z*	
[M+MeCN+H] 41.0270 *m/z*	

Custom database search

Database file: AldamaDB or AsteraceaeDB from DNP

Field order: ID, *m/z*, retention time (min), identity, formula

m/z tolerance: 0.002 or 2.5 ppm

Export/Import – Export to CSV file

APÊNDICE B - Parâmetros 2 do *MZmine 2.9.*

Raw data import

Raw data filtering: algoritmos Savitzky-Golay ou Mean Filter, ambos selecionados por predifinição (*by default*).

Peak detection

Mass detection: algoritmos Wavelet transform → Noise level: 1.0E5; Scale level: 5; Wavelet window: 20%; MS level: 1; Centroid → Noise level: 1.0E5; MS level: 1.

FTMS shoulder peaks filter (Apenas para o Orbitrap):

Mass resolution: 70,000; Peak model function: Lorentzian extended.

Chromatogram builder:

- TOF: Min time span (min): 0.2; Min heigh: 1.0E7; *m/z* tolerance: 0.06 or 4.0 ppm
- Orbitrap: Min time span (min): 0.1; Min heigh: 1.0E7; *m/z* tolerance: 0.002 or 2.5 ppm

Smoothing: 11

Chromatogram deconvolution

- TOF e Orbitrap: algoritmo Baseline Cut-off → Min. Peak height: 1.0E6; Peak duration range (min): 0.0 a 1.0; Baseline level: 1.0E6
- TOF modo positivo e Orbitrap: algorithm Local minimum search → Chromatographic threshold: 10%; Search minimum in RT range (min): 0.1; Minimum relative height: 5%; Minimum absolute height: 1.3E6; Min ratio of peak top/edge: 1; Peak duration range (min): 0.0-1.0
- TOF modo negativo: algoritmo Noise Amplitude → Min. Peak heig: 1.0E6; Peak duration range (min): 0.0 a 1.9; Amplitude of noise: 1.0E6

Isotopes – Isotopic peaks grouper

- TOF: *m/z* tolerance: 0.06 or 4.0 ppm; Retention time tolerance: 0.2 absolute; Maximum charge: 2; Representative isotope: Most intense
- Orbitrap: *m/z* tolerance: 0.002 or 2.5 ppm; Retention time tolerance: 0.1 absolute; Maximum charge: 2; Representative isotope: Most intense

Alignment

Algoritmo RANSAC

- TOF: m/z tolerance: 0.06 or 4.0 ppm; RT tolerance: 0.1; RT tolerance after correction: 10; RANSAC interation: 3000; Minimum number of points: 50%; Threshold value: 1; Linear Model: [ ]; Requiere same charge state: [x].
- Orbitrap: m/z tolerance: 0.002 or 2.5 ppm; RT tolerance: 0.1; RT tolerance after correction: 10; RANSAC interation: 3000; Minimum number of points: 50%; Threshold value: 1; Linear Model: [ ]; Requiere same charge state: [x].

Gap filling

Same RT and m/z range gap filler

- TOF: m/z tolerance: 0.06 or 4 ppm
- Orbitrap: m/z tolerance: 0.002 or 2.5 ppm

Export/Import – Export to CSV file

APÊNDICE C - Parâmetros do *MetAlign* (versão 041012).

Accurate mass mode

Mass resolution: 16,000 (TOF); 70,000 (Orbitrap)

Retention begin (Scan nr): 1

Retention end (Scan nr): 5000

Maximum amplitude: 1,000,000

Peak slope factor (x Noise): 1

Peak threshold factor (x Noise): 2

Peak threshold (Abs. Value): 5000

Average peak width at half height (Scans): 5

Scaling options: No scaling

Initial peak search criteria (*by default*):

	Scan Nr.	Max. Shift
Begin of 1st Region	0	10
End of 1st Region	4787	10
Begin of 2nd Region	0	0
End of 2nd Region	0	0

Tuning alignment options and criteria: Pre-align processing (Iterative)

Maximum shift peak per scan: 20

Max. peak section:

	1st Iteration	Last iteration
Min. factor (x Noise)	3.0	3.0
Min. Nr. of masses	4.0	3.0

Run Scaling and Alignment

APÊNDICE D - Parâmetros do *MSClust.*

Para MetAlign

Effective Peaks: 50

Peak Width, scans or r.time: 10

P.W. Margin Softness: 0.5

Correlation Threshold: 0.8

C.T. Margin Softness: 0.02

PD Reduction (0.0 – 1.0): 0.8

PD Reduction Softness: 0.02

Stop Criterion: 1

Para MZmine 2.9

Effective Peaks: 50

Peak Width, scans or r.time: 2

P.W. Margin Softness: 0.5

Correlation Threshold: 0.5

C.T. Margin Softness: 0.01

PD Reduction (0.0 – 1.0): 0.5

PD Reduction Softness: 0.01

Stop Criterion: 1

APÊNDICE E - Locais de coleta de *Aldama* La Llave.

Espécie	Local de coleta
Aldama anchusifolia	Eldorado do Sul (RS) 30°01’12.3” S; 51°19’10.1” W
Aldama arenaria	Itirapina (SP) 22°13’ S; 47°54’ W
Aldama aspillioides	Ponta Grossa (PR) 25°07’43.8” S; 50°01’57.6” W
Aldama bakeriana	Oliveira (MG) 20°46’28.1” S; 44°46’52.5” W
Aldama bracteata	Curvelo (MG) 18°35’50.5” S; 46°57’12.9” W
Aldama discolor	Sacramento (MG) 20°00’26.1 S; 47°24’30.3” W
Aldama filifolia	Alto Paraíso de Goiás (GO) 13°58’16.1” S; 47°29’56.7” W
Aldama gardneri	Alto Paraíso de Goiás (GO) 13°55’36.1” S; 47°25’51.3” W
Aldama goyazii	Ituiutaba (MG) 18°51’03.6” S; 49°28’32.3” W
Aldama grandiflora	São Roque de Minas (MG) 20°14’48.7” S; 46°26’24.8” W
Aldama kunthiana	Niquelândia (GO) 14°25’23.0” S; 48°26’12.0” W
Aldama linearifolia	Ponta Porã (MS) 22°16’38.5” S; 55°40’28.8” W
Aldama megapotamica	Butiá (RS) 30°06’31.7” S; 51°48’18.7” W
Aldama nudibasilaris	Oliveira (MG) 20°39’39.8” S; 44°44’14.8” W
Aldama nudicaulis	Pinheiro Machado (RS) 31°31’04.4” S; 53°30’48.6” W
Aldama oblongifolia	Piranhas (GO) 16°39’14.4” S; 51°42’29.4” W
Aldama pilosa	Campestre da Serra (RS) 28°46’49.7” S; 51°05’45.2” W
Aldama robusta	Altinópolis (SP) 21°01’49.9” S; 47°18’25.4” W
Aldama robusta var. *oxylepis*	Pedregulho (SP) 20°15’27.0”S; 47°28’39.00”W
Aldama rubra	Pirassununga (SP) 21°57’4.3” S; 47°22’47.4” W
Aldama santacatarinensis	São José dos Ausentes (RS)
Aldama squalida	Piranhas (GO) 16°35’44.5” S; 51°55’57.7” W
(Aldama tenuifolia	São Sebastião do Paraíso (MG) 20°58’53.0” S; 46°58’52.0” W
Aldama trichophylla	Palmeira (PR) 25°22’53.40” S; 49°48’21.00”W
Aldama tuberosa	Quaraí (RS) 30°11’22.5” S; 56°29’39.4” W
Aldama tuberosa var. *guaranitica*	Alegrete (RS) 29°39’42.24”S; 55°23’48.60”W

APÊNDICE F - Outras espécies e gêneros.

Achillea millefolium Ledeb. (mil-folhas)

Achyrocline satureioides (Lam.) DC. (marcela)

Acmella oleraceae (L.) R.K. Jansen (jambú)

Ageratum conyzoides L. (mentrasto)

Anteremanthus hatschbachii H. Rob.

Arctium lappa L. (bardana)

Arnica montana L.

Artemisia absinthium L. (losna)

Artemisia annua L.

Baccharis dracunculifolia D.C. (alecrim do campo)

Baccharis trimera (Less.) DC. (carqueja)

Bidens pilosa L. (picão)

Calea cuneifolia DC.

Calendula officinalis L. (calêndula)

Chronopappus bifrons (DC. ex Pers.) DC

Cichorium intybus L. (chicória)

Cynara scolymus L. (alcachofra)

Dasyphyllum brasiliense var. latifolium (D.Don) Cabrera (espinho-agulha)

Echinacea purpurea (L.) Moench (equinácea)

Emilia sonchifolia L. DC (emília)

Eremanthus polycephalus (DC.) MacLeish

Gymnanthemum amygdalinum (Delile) Sch. Bip. ex Walp (boldo da Bahia)

Helianthus annuus L. (girassol)

Heterocoma gracilis Loeuille, J. N. Nakaj. & Semir

Lactuca sativa L. (alface)

Lychnophora diamantinana Coile & S.B. Jones

Lychnophora ericoides Mart. (arnica-da-serra)*

Lychnophora tomentosa (Mart. ex DC.) Sch. Bip.

Matricaria chamomilla L. (camomila)

Mikania glomerata Sprengl. (guaco)

Mikania hirsutissima DC.

Mikania laevigata Schultz Bip. ex Baker (guaco(-)cheiroso)

Minasia scapigera H. Rob.

Piptolepis monticola Loeuille

Prestelia eriopus Sch. Bip.

Puchea guitoc D.C.

Smallanthus sonchifolius (Poepp. & Endl.) H. Robinson (yacón)

Solidago microglossa DC. (arnica brasileira)

Sonchus oleraceus L. (serralia)

Sphagneticola trilobata (L.) Pruskei (mal-me-quer)

Stevia rebaudiana (Bertoni) Bertoni (estévia)

Tridax procumbens L.

Tanacetum parthenium L. (feverfew)

Tanacetum vulgaris L.

Taraxacum officinale Weber ex FH Wigg. (dente-de-leão)

Tithonia diversifolia (Hemsl.) A. Gray (margaridão)

Vernonia herbaceae (Vell.) Rusby

Vernonia platensis (Spreng.) Less.

Vernonia polyanthes Less. (assa-peixe)

Vernonia rubriramea Mart. Ex DC.